Gauge Theory and Related Topics
by Nestor Azziz

Lectures Delivered to the New College of Florida
Sarasota, FL

Note: These notes were typed by Annabel Azziz,
Andrew Fiorello, and Nicholas Wojtalewicz.

This book will be uploaded to the web as a free download. If any mistakes should be found, the author would very much appreciate being told about them at nestorazziz@gmail.com.

The author would like to thank Professor Marcelo Ubriaco for encouraging him to publish these lectures.

Contents

0.1 Lecture 1 8
0.1.1 Frenet Triad 8
Example: Helix of Unit Velocity 8
Oscular Plane and Geodesic 10
0.1.2 Connecting Functions 14
0.2 Lecture 2 16
0.2.1 Metric Tensor 16
Example 1 17
0.2.2 Euclidian Space 17
Simple Examples: Vector as Tensor 18
0.2.3 Tensors 21
0.2.4 Operations with Elements 23
Addition 23
Changing from Covariant to Contravariant and Vice-Versa 24
0.2.5 Tensor Density 25
0.2.6 Notes on the Jacobian 27
Differential Volume 30
0.3 Lecture 3 32
0.3.1 Absolute Tensor Derivative 32
Absolute Derivative of the Metric Tensor, g_{ij} . 34
Christoffel Symbols as Functions of the Derivative of the Metric Tensor . . . . 35 [redacted]
Absolute Derivative of a Tensor of the Second Order (Two Indices), Tensor t_{ij}: . . . 36

Note on the Absolute Derivative of a Second Order Tensor t_{ij}: 36
Summary of the Absolute Tensor Derivative of the Previous Examples 38
NOTE: Advanced Conclusions on Parallel Transport as it will be Discussed in the Next Section . . . . 40 [redacted]
Summary: Relationship Between the Christoffel Symbols and the Metric Tensor . 41 [redacted]
Examples of Christoffel Values for Curvilinear Coordinates 42
Christoffel Symbols in Polar Coordinates and Their Relationship to the Centripetal Force . . . . 43 [redacted]
0.4 Lecture 4 45
0.4.1 Riemann Space 45
Example of a Geodesic and Parallel Transport 48
Curvature of the Space and Riemann Space . 49
Space With Normal Coordinates 54
Ricci Tensor 54
Bianchi Identity 55
0.5 Lecture 5 60
0.5.1 Weyl Space 60
Variation of the Length of a Vector as it Moves Parallel to Itself: Weyl Geometry . . 60
Transport of a scalar l around a closed curve . 60
Generalization of the Ricci Theorem 64
Invariance of the Angle Between Two Vectors as They Move Parallel to Themselves 64
Relationship Between Weyl $\bar{\Gamma}$ and the Christoffel Symbols . . . . 65 [redacted]
Curvature and Generalized Christoffel Symbols 66
0.6 Lecture 6 67
0.6.1 Gauge Theory 67

Introducing Different Gauges, or "Etalon," in the Same Space 67
Invariance of the Absolute Derivative of ξ^i . . 71
Invariance of Curvature as Gauge Changes . . 72
Appendix 74
0.7 Lecture 7 80
0.7.1 Analytic Dynamics 80
Variational Calculus and the Euler Differential Equation 80
Application of the Euler Differential Equation: Geodesic from the Variational Point of View 83
Geodesic and absolute derivatives 86
Geodesic: Summary 87
D'Alembert Principle 89
Problems Using the D'Alembert Principle and the Euler-Lagrange Equation 92
Euler–Lagrange–Poisson–Jacobi–Hamilton–Canonical Equations 96
Lagrange–Poisson–Jacobi–Maupertuis 98
Maxwell Equations and Their Invariance Under Gauge Change 101
Tensor Form of the Maxwell Equations 103
Review of Relativity 106
Relativity According to Minkowsky 112
Relativistic Dynamics: Application of Tensor Theory to Special Relativity 123
Relativistic Forces 127
Relativistic Lagrangian and the Lorentz Force 135
Hamiltonian for a Charged Particle 139
Note on Maupertuis and Fermat 141
Canonical Equations of Hamilton–Poisson Brackets and Jacobi Algebra . . . . 143 [redacted]
Poisson Brackets 144

Mathematicians That Contributed to Analytic Mechanics 146

0.8 Lecture 8 147

0.8.1 Matrices, Operators, and Elementary Group Theory 147

The Adjoint Matrix (a^+) 149

Examples of Linear Operators 152

Some Geometric Uses of Matrices 152

Group Theory: A Review 154

0.8.2 Review of Lie Algebra 159

Infinitesimal Transformations 159

Lorentz Generators 167

Lorentz Generators and Lie Algebra 174

0.9 Lecture 9: Fundamentals of Today's Cosmology . . . 178

0.9.1 Differential Operators in Curvilinear Coordinates 178

The Gradient ∇ of the Function Φ 181

The Rotational $\nabla\times$ of the Vector $\vec{F}$ 181

The Divergence $\nabla\cdot$ of the vector $\vec{F}$ 181

The Laplacian Δ of the Function Φ 183

0.9.2 Potential Theory and the Hamilton-Jacobi Equation 184

Relationship of Potential Theory to the Poisson Bracket 185

Differential Equations for Potential Theory . . 186

Some Typical Potentials in Classic Physics and the Corresponding Differential Equations 187

Examples 188

0.9.3 Simplifications Regarding the the Space-time "Length" Invariant, the Geodesic Equation, the Red Shift, Black Holes, and the Einstein Approximation 190

1.) The Space-time "Length" Invariant 190

2.) The Geodesic Equation 190
3.) The Red Shift Due to Gravitational Effects and the Expansion of the Universe 193
4.) Classic Visualization of Black Holes 196
Simplified Black Hole Derivations Using Equations (0.9.10) and (0.9.13) . . . . . . 197 [redacted]
5.) The Einstein Relationship in Different Approximations 199
Subapproximation for the Einstein Relationship 200
Simplifications for the Ricci Tensor and the Gravitational Einstein Equations in a New Approximation 201
0.9.4 Calculation of the Christoffel Symbols Under Some Simplifications 203
The Schwarzchild Metric Tensor 205
0.9.5 The Hubble Law 210
0.9.6 Expansion of the Universe Under the Friedmann Fundamental Metric and a Tensor Force Acting on the Material of the Galaxy 215
A Particular Case of the Tensor Force $T^{\mu\nu}$ . . 222

0.1 Lecture 1

0.1.1 Frenet Triad

Parametric Curves are functions which are defined in terms of useful and related quantities, often in *time* or *arc length.* Before going on much further, it is necessary to define a few expressions:

- The velocity (or *tangent vector*) can be thought of as the change in position with respect to arc length, or $\vec{T} = \dfrac{\Delta \vec{f}}{\Delta s} = \vec{f}'(s)$. We also want to define the *unit tangent vector* such that $\hat{\mathbf{T}} = \dfrac{\vec{T}}{|T|}$, where its magnitude is always equal to 1.

- From the tangent vector we can define the normal vector as $\frac{\vec{T}'}{|\vec{T}'|}$, or in other terms, $\vec{T}' = k\vec{N}$. k here is called the curvature; to be specific it is the *extrinsic* curvature.

- The Binormal vector is defined as $\vec{B} = \vec{T} \times \vec{N}$. The derivative of the this vector can be related to the normal vector N by $\vec{B}' = -\tau\vec{N}$ using a quantity called *torsion*, represented by τ.

- Since $\vec{N} = \vec{B} \times \vec{T}$,

  $$\vec{N}' = \vec{B}' \times \vec{T} + \vec{B} \times \vec{T}' = -\tau(\vec{N} \times \vec{T}) + k(\vec{B} \times \vec{N}) = \tau\vec{B} - k\vec{T}$$

  Together, these quantities form the *Frenet Formulas*:

  $$\begin{bmatrix} \vec{T}' \\ \vec{N}' \\ \vec{B}' \end{bmatrix} = \begin{bmatrix} 0 & k & 0 \\ -k & 0 & \tau \\ 0 & -\tau & 0 \end{bmatrix} \begin{bmatrix} \vec{T} \\ \vec{N} \\ \vec{B} \end{bmatrix}$$

Example: Helix of Unit Velocity

Let us look at the function

$$\vec{f}(s) = a\cos\left(\frac{s}{c}\right)\hat{\mathbf{i}} + a\sin\left(\frac{s}{c}\right)\hat{\mathbf{j}} + \frac{b}{c}s\hat{\mathbf{k}} \ ,$$

where a, b and c are arbitrary constants and s is the length of the curve itself. It could equivalently be written by its components

$$x(s) = a\cos\left(\frac{s}{c}\right) \quad , \quad y(s) = a\sin\left(\frac{s}{c}\right) \quad \text{and} \quad z(s) = \frac{b}{c}s \ .$$

Plotted, this curve gives us a circular spiral that rises or falls depending on the ratio of b to c. Note that if $b = 0$ then we have a two-dimensional curve $(f'(s))$ that traces out a circle on the xy-plane. The tangent vector can be found to be

$$\vec{T}(s) = \left(-\frac{a}{c}\right)\sin\left(\frac{s}{c}\right)\hat{\mathbf{i}} + \left(\frac{a}{c}\right)\cos\left(\frac{s}{c}\right)\hat{\mathbf{j}} + \left(\frac{b}{c}\right)\hat{\mathbf{k}} \ .$$

By analyzing the tangent vector, we can guess the following orthogonal vector in the (x, y) plane:

$$\vec{N}(s) = \left(\frac{c}{a}\right)\cos\left(\frac{s}{c}\right)\hat{\mathbf{i}} + \left(\frac{c}{a}\right)\sin\left(\frac{s}{c}\right)\hat{\mathbf{j}} + 0\hat{\mathbf{k}}$$

(observe that we chose $|N| = \frac{c}{a}$). From this, we can find the derivative to be

$$\vec{N}'(s) = \left(-\frac{1}{a}\right)\sin\left(\frac{s}{c}\right)\hat{\mathbf{i}} + \left(\frac{1}{a}\right)\cos\left(\frac{s}{c}\right)\hat{\mathbf{j}} + 0\hat{\mathbf{k}} \ .$$

With the tangent and normal vectors, we can also find the binormal and its derivative:

$$\vec{B}(s) = \vec{T} \times \vec{N} = \left(-\frac{b}{a}\right)\sin\left(\frac{s}{c}\right)\hat{\mathbf{i}} + \left(\frac{b}{a}\right)\cos\left(\frac{s}{c}\right)\hat{\mathbf{j}} - 1\hat{\mathbf{k}} \ , \text{ and}$$

$$\vec{B}'(s) = \left(-\frac{b}{ca}\right)\cos\left(\frac{s}{c}\right)\hat{\mathbf{i}} - \left(\frac{b}{ca}\right)\sin\left(\frac{s}{c}\right)\hat{\mathbf{j}} + 0\hat{\mathbf{k}} \ .$$

Comparing with the equation for $\vec{N}$,

$$\vec{B}' = -\frac{b}{c^2} \to \tau = \frac{b}{c^2} = \frac{b}{a^2 + b^2} \ .$$

If $b = 0$, then there is no torsion $(\tau = 0)$.

Observe that if $\vec{T}$ is a unit vector, then $a^2 + b^2 = c^2$.

From the equations for $\vec{T}'$ and $\vec{N}$,

$$\vec{T}' = -\frac{a^2}{c^3} N \therefore k = -\frac{a^2}{c^3} \ .$$

The parametric function of s can be Taylor expanded with a differential Δs. This expansion immediately reveals the quantities we defined above:

$$f(s) = f(0) + \underbrace{f'(0)}_{T_0} \Delta s + \underbrace{f''(0)}_{k_0 N_0} \frac{(\Delta s)^2}{2} + \underbrace{f'''(0)}_{N_0{}' k_0} \frac{(\Delta s)^3}{6} + \ldots$$

and, if we only look at a small portion of the curve, we can assume that $k \cong$ constant. Then,

$$f(s) = f(0) + T_0 \Delta s + k_0 N_0 \frac{(\Delta s)^2}{2} + k_0(\tau_0 B_0 - k_0 T_0) \frac{(\Delta s)^3}{6} + \ldots$$

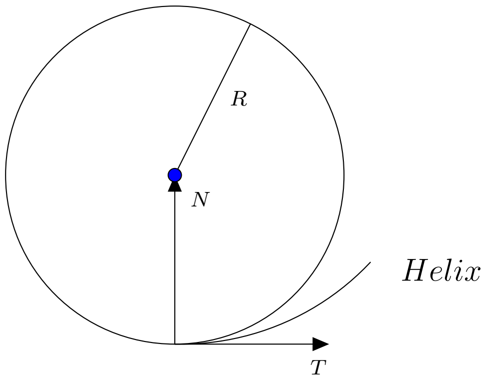
In the last term we neglect a second order one $(k_0^2 T_0^2)$, and the final form of $f(s)$ for small increments is:

$$f(s) = f(0) + T_0 \Delta s + k_0 N_0 \frac{(\Delta s)^2}{2} + k_0 \tau_0 B_0 \frac{(\Delta s)^3}{6} \qquad (0.1.1)$$

Oscular Plane and Geodesic

The helix is an interesting case that can be used to introduce the concepts of the oscular plane and the geodesic.

In the case discussed above, the oscular plane is the one formed by the tangent $\vec{T}$ and the normal $\vec{N}$. The vector $\vec{B}$ is perpendicular to this plane. In this plane, we may draw a circle that is at the point in consideration to the tangent $\vec{T}$, and that also contains the differential segment of the helix that encloses the point. Thus, the circle has the same curvature R as the helix $(R = k)$.



Next, we will show that if the helix is the curve that joins two points A and B on the surface of a cylinder, then it is the shortest distance between the two points, just like a straight line is the shortest distance between two points in a plane. The curve that fulfills these properties is called the *geodesic*.

Assume a point moving on a cylinder from A to B. We will show that if the distance between A and B as measured by the length of the curve is the shortest possible, then it is a helix. If we take s as the parameter on the curve, then the coordinates of a point on the curve are

$$x = a\cos(s)$$
$$y = a\sin(s)$$
$$z = z$$

or

$$\mathrm{d}x = (-a\sin s)\mathrm{d}s$$
$$\mathrm{d}y = (a\cos s)\mathrm{d}s$$
$$\mathrm{d}z = \mathrm{d}z\ .$$

The function $z(s)$ is to be determined. The differential length $\mathrm{d}l$ is such that

$$(\mathrm{d}l)^2 = (\mathrm{d}x)^2 + (\mathrm{d}y)^2 + (\mathrm{d}z)^2 = (a\mathrm{d}s)^2 + (\mathrm{d}z)^2\ .$$

Thus, we want to minimize the following integral:

$$I = \int_A^B \mathrm{d}l = \int_A^B \sqrt{a^2(\mathrm{d}s)^2 + (\mathrm{d}z)^2} = \int_A^B \mathrm{d}s\sqrt{a^2 + z'^2}$$

where $z' = \frac{\mathrm{d}z}{\mathrm{d}s}$. The problem of minimizing an integral belongs to the subject of Variational Calculus. This branch of calculus is discussed in Lecture 7: Review of Analytic Dynamics. It may be presented in the following way: if the function $f(s, y(s), y'(s))$ is to be integrated between the point A and B while s is a parameter (it could be, for instance, length or time), then the integral

$$I = \int_A^B f(s, y(s), y'(s))\ \mathrm{d}s$$

will have a minimum or maximum when $y(s)$ fulfills the following differential equation (named after Euler, the mathematician who found the solution in 1700):

$$\frac{\partial f}{\partial y} - \frac{\mathrm{d}}{\mathrm{d}s}(\frac{\partial f}{\partial y'}) = 0\ .$$

In our case,

$$f = \sqrt{a^2 + z'^2}$$

and z is the function y and thus $z' = y' = \frac{\mathrm{d}z}{\mathrm{d}s}$.

The Euler equation says

$$\frac{\partial^2}{\partial s}\frac{f(s)}{\partial z'} = 0$$

or

$$\frac{\mathrm{d}}{\mathrm{d}s}\left(\frac{1}{2}\frac{2z'}{\sqrt{a^2 + z'^2}}\right) = 0\ .$$

Thus,

$$\frac{z'}{\sqrt{a^2 + b^2}} = c$$

where c is any constant. And so,

$$z'^2 = \frac{a^2c^2}{1 - c^2} = D^2$$

or

$$z = Ds \ .$$

where D is a constant; so our curve with the minimum length between two points is a helix. Since D could be any constant, there are an infinite number of helices.

The fact that the helix is a geodesic is easily seen if the cylinder is unfolded and converted into a plane with width $2\pi r$. The helical coils become straight lines, and obviously the shortest path between two points A and B is a segment of a coil. We note that on the cylinder, as mentioned before, we may create an infinite number of geodesics, each characterized by a slope α. The slope of the helix passing through two points A and B is predetermined by those two points.

Another interesting example that helps to visualize the meanings of the oscular plane and the geodesic is the case of a thread on a sphere; the only force acting on the thread is the normal reaction from the sphere. The normal reaction passes through the center of the sphere at any point on the thread. So, the plane containing the normal is the oscular plane, which cuts the sphere into a grand circle. The grand circle through the equator and prime meridian is the geodesic of the sphere. If the arc of that circle is greater than 180 degrees, then the distance between two point A and B on this arc is not the shortest distance, because the shortest distance is the supplement arc. So, in this case we have a geodesic which is not necessarily the shortest distance between two points.

0.1.2 Connecting Functions

As in the case of Frenet frames, when the frame moves, the derivatives of $\vec{T}$, $\vec{N}$, and $\vec{B}$ give the intrinsic properties of the curve – its curvature (k) and its torsion (τ). These derivatives with respect to the curve parameter are called "covariant derivatives."

It's interesting to note that those derivatives ($\vec{T}'$, $\vec{N}'$, and $\vec{B}'$) can be expressed as functions of the frame itself ($\vec{T}$, $\vec{N}$, and $\vec{B}$). We can think of these spatially as the variation from some point P to some point $P + \delta P$.

We can define this variation a bit more formally if we start in some general frame: $\vec{e}_1, \vec{e}_2, \ldots, \vec{e}_n$ such that $\vec{e}_i \cdot \vec{e}_j = g_{ij}$. In this frame, we can define the variation of the vector $\mathrm{d}\vec{e}_i$ as

$$\mathrm{d}\vec{e}_i = W_{ij} \cdot \vec{e}_j \ .$$

Here, W_{ij} is a function that depends on the point P in our generalized coordinate space $(q_1, q_2, \ldots, q_n)$ where the frame $\vec{e}_1, \vec{e}_2, \ldots, \vec{e}_n$ is located. It is closely related to the matrix form for Frenet's formulas previously shown: in the simple case of the Frenet frame, the derivatives of $\vec{T}$, $\vec{N}$, and $\vec{B}$ with respect to the parameter s which varies along the tangent $\vec{T}$ are the covariant derivatives. As such, $\vec{e}_1 \equiv \vec{T}$, $\vec{e}_2 \equiv \vec{N}$, $\vec{e}_3 \equiv \vec{B}$, and

$$\mathrm{d}\vec{e}_1 = W_{ij}\vec{e}_j \ .$$

The matrix W_{ij} is

$$W_{ij} \propto \begin{pmatrix} T' \\ N' \\ B' \end{pmatrix} \begin{pmatrix} 0 & k & 0 \\ -k & 0 & \tau \\ 0 & -\tau & 0 \end{pmatrix}$$

Considering the special case of an orthogonal frame ($\vec{e}_i \cdot \vec{e}_j = \delta_{ij}$, where δ_{ij} is the Kronecker-Delta Function), we can use the product rule to expand that relation to $(\mathrm{d}\vec{e}_i)\vec{e}_j + \vec{e}_i(\mathrm{d}\vec{e}_j) = 0$. Using this, we can determine

$$W_{ij} = -W_{ji} \text{ , and } W_{kk} = 0 \ .$$

W_{ij} is a *connection* function which depends on point P and it expresses the differential variation of $\vec{e}_i$ in the direction of $\vec{e}_j$. If we act under the assumption that W_{ij} is a linear combination of coordinate differentials $\mathrm{d}q_1, \mathrm{d}q_2, \ldots, \mathrm{d}q_i, \ldots, \mathrm{d}q_j$ then it can be rewritten as

$$W_{ij} = \mathrm{d}(\gamma_{ij}) = \Gamma_{ij}^{k} \cdot \mathrm{d}q^{k} \ ,$$

where the coefficient Γ_{ij}^{k} was first introduced by (and is named after) E.B. Christoffel in 1869.

0.2 Lecture 2

0.2.1 Metric Tensor

A point M is determined in the Euclidian space by its coordinates with respect to a reference point 0. The vector $\overrightarrow{OM}$ in rectilinear coordinates is fixed by the n coordinates $M(x_1, x_2, \ldots, x_n)$.

The idea is to find a set of curvilinear coordinates with center in M that allows us to describe the geometry of the space around M. Let us call $y^1, y^2, \ldots, y^n$ these coordinates. The n y^i coordinates are called "normal" or "natural" coordinates at M. At $\overrightarrow{M} + \delta\overrightarrow{M}$ this system may be different.

As M moves by δM it generates a set of n vectors $\vec{e}_1$, $\vec{e}_2, \ldots,$ $\vec{e}_n$, such that

$$\vec{e}_i = \frac{\partial \overrightarrow{M}}{\partial \vec{y}^i}.$$

When introducing a metric g_{ij} in the space, the "surface" where M exists becomes a "geometric surface". The differential length $\mathrm{d}s^2 = (\delta\overrightarrow{M})^2$ is given by

$$\mathrm{d}s^2 = g_{ij}\mathrm{d}y^i\mathrm{d}y^j \ ,$$

where g_{ij} now may depend on the coordinates y^is.

If the curve described by M is parametrized with parameter t, then the length between two points a and b is

$$\delta M = \int_a^b \sqrt{g_{ij}\,\frac{\mathrm{d}y^i}{\mathrm{d}t}\,\frac{\mathrm{d}y^j}{\mathrm{d}t}}\ \mathrm{d}t \ ,$$

and the volume $\mathrm{d}v$ formed by $\vec{e}_i\mathrm{d}y^i$ with origin at M is

$$\mathrm{d}V = \sqrt{|g|}\ \mathrm{d}y^1, \ldots, \mathrm{d}y^n \ .$$

Example 1

The curvilinear coordinates at M are the spherical one with the vectors $\vec{e}_r$ along the $\overrightarrow{OM}$ direction, $\vec{e}_\psi$ on the parallel around the x^3 axis and $\vec{e}_\theta$ in the meridian passing by M. Thus, if $\hat{r}, \hat{\psi}$ and $\hat{\theta}$ are the unit vectors in these three orthogonal directions, then

$$\vec{e}_r = \hat{r} \qquad \vec{e}_\psi = r\sin(\theta)\vec{\hat{\psi}} \qquad \vec{e}_\theta = r\hat{\theta}\ .$$

The matrix element is then associated with the differential length $\mathrm{d}s^2 = \mathrm{d}M \times \mathrm{d}M$ by its Pythagoric form

$$\mathrm{d}s^2 = (\mathrm{d}r)^2 + (r\sin\theta)^2(\mathrm{d}\psi)^2 + r^2(\mathrm{d}\theta)^2\ .$$

So

$$\mathrm{M} = \begin{array}{c} \\ \hat{r} \\ \hat{\psi} \\ \hat{\theta} \end{array} \begin{array}{c} \begin{array}{ccc} \hat{r} & \hat{\psi} & \hat{\theta} \end{array} \\ \begin{pmatrix} 1 & 0 & 0 \\ 0 & (r\sin\theta)^2 & 0 \\ 0 & 0 & r^2 \end{pmatrix} \end{array}\ .$$

The change in variables from the rectilinear coordinates to the curvilinear ones and viceversa are:

$$(x_1, x_2, x_3) \longrightarrow (y^1, y^2, y^3)$$

$$r = \sqrt{x_1^2 + x_2^2 + x_3^2} \qquad \psi = \tan^{-1}\left(\frac{x_2}{x_1}\right) \qquad \theta = \tan^{-1}\left(\frac{\sqrt{x_1^2 + x_2^2}}{x_3}\right)$$

$$(y^1, y^2, y^3) \longrightarrow (x_1, x_2, x_3)$$

$$x = r\sin(\theta)\cos(\psi) \qquad y = r\sin(\theta)\sin(\psi) \qquad z = r\cos(\theta)$$

(Where the y^i are understood to be, in this context, r, θ, ψ).

0.2.2 Euclidian Space

Assume a Euclidian Space with coordinates $p^1, p^2, \ldots, p^n$. A point $M(p^1, \ldots, p^n)$ in the space moves in the direction of p^i generating a vector

$$\vec{e}_i = \frac{\partial M}{\partial p^i}\ .$$

Thus the differential of M equal dM is a vector

$$\mathrm{d}\overrightarrow{M} = \overrightarrow{e}_i \mathrm{d}p^i \ .$$

(The summation on the repeated indices is understood.)

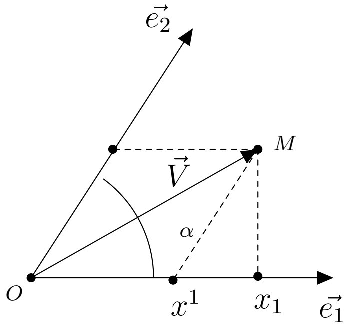
If the motion of M is on a curve, the element ds of the curve is given by

$$\mathrm{d}^2 s = \mathrm{d}\overrightarrow{M} \cdot \mathrm{d}\overrightarrow{M} = \overrightarrow{e}_i \cdot \overrightarrow{e}_j \mathrm{d}p^i \mathrm{d}p^j \ .$$

The scalar product is called g_{ij}, the metric tensor of the space. That is,

$$g_{ij} = \overrightarrow{e}_i \cdot \overrightarrow{e}_j \ .$$

Simple Examples: Vector as Tensor



Let us assume a plane in the Euclidian space. The axes are given by the unit vectors $\overrightarrow{e}_1$ and $\overrightarrow{e}_2$. A vector $\overrightarrow{V}$ has components x^1 and x^2. So

$$\overrightarrow{V} = x^1 \overrightarrow{e}_1 + x^2 \overrightarrow{e}_2 \ .$$

The components x^1 and x^2 are the ***contravariant*** components of $\overrightarrow{V}$. The projection of $\overrightarrow{V} \equiv \overrightarrow{OM}$ on $\overrightarrow{e}_1$ is $\overrightarrow{V} \cdot \overrightarrow{e}_1 = x_1$ and the projection of $\vec{e}_2$ is $\vec{V} \cdot \vec{e}_2 = x_2$. These are called the ***covariant*** components of $\vec{V}$.

$$\begin{aligned}\vec{V} \cdot \vec{e}_1 &= x^1(\vec{e}_1 \cdot \vec{e}_2) + x^2(\vec{e}_2 \cdot \vec{e}_1) \\ &= x_1 = x^1 g_{11} + x^2 g_{12} = x^1 + \cos(\alpha) x^2 \ ,\end{aligned}$$

where α is the angle between the $\vec{e}_1$ and $\vec{e}_2$. This equality may be verified geometrically by observing that the two shaded areas are equal.

Let us show in a simple form that the scalar product of two vectors $\vec{V}$ and $\vec{W}$ may be written as

$$\vec{V} \cdot \vec{W} = V^i W_i \quad \text{or} \quad \vec{V} \cdot \vec{W} = V^i W^i g_{ij} \ .$$

In fact,

$$V_1 = \vec{V} \cdot \vec{e}_1 \quad , \quad V_2 = \vec{V} \cdot \vec{e}_2 \quad , \quad \vec{V} = V^1 \vec{e}_1 + V^2 \vec{e}_2$$

and similarly for $\vec{W}$. Thus,

$$\begin{aligned}V^1 W_1 + V^2 W_2 &= V^1(\vec{W} \cdot \vec{e}_1) + V^2(\vec{W} \cdot \vec{e}_2) \\ &= V^1(W^1 \vec{e}_1 + W^2 \vec{e}_2) \cdot \vec{e}_1 + \ldots \\ &= V^1 W^1 g_{11} + V^1 W^2 g_{12} + V^2 V^1 g_{21} + V^2 W^2 g_{22} \ .\end{aligned}$$

In a matrix form, the transformation from contravariant to covariant of the vector $\overrightarrow{OM}$ of the figure could be written as :

$$\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} g_{11} & g_{12} \\ g_{21} & g_{22} \end{pmatrix} \begin{pmatrix} x^1 \\ x^2 \end{pmatrix} = \begin{pmatrix} 1 & \cos\alpha \\ \cos\alpha & 1 \end{pmatrix} \begin{pmatrix} x^1 \\ x^2 \end{pmatrix} \ .$$

The inverse of the matrix (g_{ij}) is (g^{ij}) such that

$$g^{ij} = \frac{\text{cofactor of } g_{ij}}{\det(g)} \text{ or } g^{-1} = \begin{pmatrix} \frac{1}{D} & -\frac{\cos\alpha}{D} \\ -\frac{\cos\alpha}{D} & \frac{1}{D} \end{pmatrix} .$$

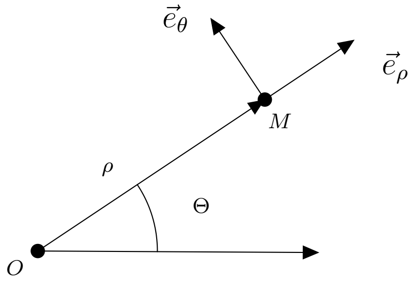
Then

$$x^i = g^{ij} x_j \ ,$$

which is a simple case of raising indices by means of the fundamental tensor. In the Euclidean space where $\vec{e}_i \cdot \vec{e}_j = \vec{e}_j \cdot \vec{e}_i$, g_{ij} is symmetric. Observe that the determinants of g^{ij} and g_{ij} are equal.

Let us see the value of g_{ij} in other curvilinear coordinates.

1. Polar coordinates



The point M is given by

$$(\rho, \theta) : \mathrm{d}s^2 = \mathrm{d}\rho^2 + \rho^2 \mathrm{d}\theta^2$$
$$\vec{e}_\theta = \rho\hat{\theta} \qquad \text{and} \qquad \vec{e}_\rho = \hat{\rho}$$

where $\hat{\theta}$ and $\hat{\rho}$ are the unit vectors. So

$$g_{\theta\theta} = \rho\hat{\theta} \cdot \rho\hat{\theta} = \rho^2 \qquad g_{\rho\rho} = \hat{\rho} \cdot \hat{\rho} = 1$$
$$g_{\theta\rho} = g_{\rho\theta} = 0$$

$$g = \begin{pmatrix} 1 & 0 \\ 0 & \rho^2 \end{pmatrix}$$

2. Spherical Coordinates.

   Then $\vec{e}_\theta = \rho\hat{\theta}$, $\vec{e}_\rho = \hat{\rho}$, $\vec{e}_\psi = \rho \sin\theta\hat{\psi}$

   where $\hat{\theta}$, $\hat{\rho}$ and $\hat{\psi}$ are the orthogonal unit vectors.

$$g_{\theta\theta} = \vec{e}_\theta \cdot \vec{e}_\theta = \rho^2 \qquad g_{\rho\rho} = 1 \qquad g_{\psi\psi} = \rho^2 \sin^2\theta$$

   and all the others are zero. Thus,

$$g = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \rho^2 & 0 \\ 0 & 0 & \rho^2 \sin^2\theta \end{pmatrix}$$

0.2.3 Tensors

Let us change the coordinates such that the space will be described by a new set of n coordinates $q^1, \dots q^n$. The coordinates p^i of M are given by the following functions:

$$p^i = p^i(q^1, \dots, q^n) \ ,$$

where the index i runs from 1 to n, and the transformation is biunique and regular. That is, the Jacobian of the transformation is different from zero (see note at the end of this section). Let us call Ω a system of several functions that we call “elements” of $\Omega : E_1, E_2, \dots, E_w$. Assume that E is just a function $F(p^1, \dots p^n)$ that changes to $Q(p^1(q^1, \dots, q^n), \dots p^n(q^1, \dots q^n))$ in the change of variables.

Obviously

$$F(p^1, \dots p^n) = Q(q^1, \dots q^n)$$

So a simple substitution keeps the “element” E invariant.

Now assume that the element is the derivative of a function. So

$$E_i(p^1, p^2, \dots, p^n) = \frac{\partial F(p^1, p^2, \dots, p^n)}{\partial p^i}$$

By changing variables $E_i(p^1,\dots,p^n) \longrightarrow Q_j(q^1,\dots q^n)$ such that

$$Q_j = \frac{\partial F(p^1(q^1,\dots,q^n),\dots,p^n(q^1,\dots,q^n))}{\partial p^i}\left(\frac{\partial p^i}{\partial q^j}\right)$$

we find that a simple substitution of the variables will not preserve the element. It will give the first factor only. The second factor that is the ratio of the old variables $p^1,\dots,p^n$ to the new ones, $q^1,\dots,q^n$, is necessary to preserve the character of the element, the geometric meaning of the derivative.

Elements like Q_j are called covariant elements. We identify it by a subscript and they transform by having a factor which is the ratio of the old coordinates to the new ones, symbolically written as $\frac{\partial p}{\partial q}$.

The metric g_{ij} is also a covariant element. Since there are two indexes i and j in the transformation there will be two ratio-factors. In fact,

$$\mathrm{d}s^2 = g_{ij}\mathrm{d}p^i\mathrm{d}\rho^j = g_{ij}\frac{\partial p^i}{\partial q^r}\mathrm{d}q^r\frac{\partial p^j}{\partial q^s}\mathrm{d}q^s = \rho_{rs}\mathrm{d}q^r\mathrm{d}q^s \ ,$$

where

$$\rho_{rs} = g_{ij}\left(\frac{\partial p^i}{\partial q^r}\right)\left(\frac{\partial p^j}{\partial q^s}\right) \ .$$

So g_{ij} is a covariant with two indices (or of order 2). Observe that despite the fact that g_{ij} may be a function of the coordinates, our derivation is correct since we just expressed $\mathrm{d}p^i$ and $\mathrm{d}p^j$ as a function of the q^i's. Let us assume now that the element is a factor A^r which forms a linear combination of the derivatives of a function of $(p^1,\dots,p^n)$. That is, the linear combination:

$$A^r\frac{\partial f(p^1,\dots,p^n)}{\partial p^r},$$

which, after the change of coordinates $f(p^1,\dots,p^n) \to h(q^1,\dots,q^n)$

$$A^r\frac{\partial f(p^1,\dots,p^n)}{\partial p^r} = B^s\frac{\partial h(q^1,\dots,q^n)}{\partial q^s} = B^s\frac{\partial f(p^1,\dots,p^n)}{\partial p^r}\left(\frac{\partial p^r}{\partial q^s}\right)$$

So the new factor B^s is related to A^r by

$$B^s = A^r \frac{\partial q^s}{\partial p^r}.$$

The element A^r is said to be a contravariant tensor. It changes to B^s by the ratio of the new coordinates to the old ones. In general, if P is the element with n number of covariant indices, the corresponding Q after the transformation $p \to q$ is written as

$$P_{r_1,\ldots,r_n} = Q_{s_1,\ldots,s_n} \frac{\partial p^{s_1}}{\partial q^{r_1}} \cdots \frac{\partial p^{s_n}}{\partial q^{r_n}}.$$

And for the contravariant system

$$P^{r_1,\ldots,r_n} = Q^{s_1,\ldots,s^n} \frac{\partial q^{r_1}}{\partial p^{s_1}} \cdots \frac{\partial q^{r_n}}{\partial p^{s_n}} \ .$$

Every index r_i or s_j takes the n values of the coordinates:

$$q^1, \ldots, q^n \qquad \text{or} \qquad p^1, \ldots, p^n$$

in the second side of the equation. Repeated indexes imply addition.

0.2.4 Operations with Elements

Addition

Two covariant or contravariant elements of the same order may be summed to result in another element of the same type, with the same order,

$$S_{r_1,\ldots r_n} = A_{r_1,\ldots r_n} + B_{r_1,\ldots r_n} \ .$$

This means that (it can be easily proved) as the coordinates change $p \longrightarrow q$,

$$S_{r_1,\dots r_n}(p) \longrightarrow Q_{s_1,\dots s_n}(q)\frac{\partial p^{s_1}}{\partial q^{r_1}}\cdots\frac{\partial p^{s_m}}{\partial q^{r_n}} \ .$$

These elements may be also multiplied with a result of another element whose order is the sum of the order of the factors

$$A_{r_1,\dots r_n} \times B_{s_1,\dots s_n} = R_{r_1,\dots r_n,s_1,\dots s_n} \ .$$

For contravariant elements we have similar relationships.

It also can be shown that when we multiply elements with covariant and contravariant indices the result is an element with less indices, as in the next example:

$$A^{s_1,\dots s_n}B_{r_1,\dots r_n,s_1,\dots s_n} = R_{r_1,\dots r_n} \ .$$

Changing from Covariant to Contravariant and Vice-Versa

From the previous equation that expresses the contraction or saturation of indices we may see that the following equation is an identity:

$$g^{rs} = g^{ri}g^{sj}g_{ij} \ .$$

We observe that by definition

$$g^{\mu\nu} = \frac{\mathcal{C}^{\mu\nu}}{|g|} \ .$$

where $\mathcal{C}^{\mu\nu}$ is the cofactor of $g^{\mu\nu}$. Thus

$$|g| = g_{\mu\nu}\mathcal{C}^{\mu\nu} = g_{\mu\nu}g^{\mu\nu}|g|$$

and the following must therefore be true:

$$g_{\mu\nu}g^{\mu\nu} = 1 \ .$$

Another way of writing this is

$$g_{\mu\nu}g^{\alpha\beta} = \delta^{\alpha}_{\mu}\delta^{\beta}_{\nu}$$

where $\delta^{\alpha}_{\mu} = 1$ if $\alpha = \mu$ and is zero otherwise, and $\delta^{\beta}_{\nu} = 1$ if $\beta = \nu$ and is zero otherwise.

In general, we can raise or lower indices, i.e. go from contravariant to covariant indices and viceversa, by using the fundamental tensor g_{ij}. For instance,

$$R^{r_1,\ldots r_n} = g^{r_1 s_1} \ldots g^{r_n s_n} R_{s_1,\ldots s_n}.$$

For a contravariant element,

$$R_{r_1,\ldots r_n} = g_{r_1 s_1} \ldots g_{r_n s_n} R^{s_1,\ldots s_n}.$$

This and previous sections have been based on the works of M. Christoffel (1859) and that of M.M.G. Ricci and T. Levi-Civita (1898, *Methodes de Calcul Differential Absolu et leurs Application*). What we have called an "element" here is in today's language a "tensor."

0.2.5 Tensor Density

In a change of reference from $p \to q$, $\overrightarrow{e_\mu} \to \overrightarrow{e'}_\mu$, and $g_{ij} \longrightarrow f_{ij}$; then

$$(\mathrm{d}s)^2 = g_{ij}\mathrm{d}p^i\mathrm{d}p^j = f_{rs}\mathrm{d}q^r\mathrm{d}q^s \ ,$$

where

$$f_{rs} = g_{ij}\frac{\partial p^i}{\partial q^r}\frac{\partial p^j}{\partial q^s} \ .$$

We put this into a matrix form by introducing

$$a^i_r = \frac{\partial p^i}{\partial q^r} \text{ and } a^j_s = \frac{\partial p^j}{\partial q^s} \ ;$$

thus,

$$f_{rs} = a^i_r a^j_s g_{ij} \ .$$

Taking the determinant of the above matrix equation,

$$|f_{rs}| = |a^i_r||a^j_s||g_{ij}| = J^2|g_{ij}|$$

where J is the Jacobian of the transformation. Thus,

$$\sqrt{|f_{rs}|} = J\sqrt{|g_{ij}|} \ .$$

In the four dimensional space of Minkowsky,

$$p^1 = ct \qquad p^2 = x \qquad p^3 = y \qquad p^4 = z \ .$$

The element under coordinate transformation is related to the light path:

$$(\mathrm{d}s)^2 = c^2(\mathrm{d}t)^2 - (\mathrm{d}x)^2 - (\mathrm{d}y)^2 - (\mathrm{d}z)^2 \ .$$

The fundamental matrix in this space is then

$$g_{ij} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix} \ .$$

The element differential “volume” in 4-D space with respect to the old variables is

$$\mathrm{d}v = \mathrm{d}p^1\,\mathrm{d}p^2\,\mathrm{d}p^3\,\mathrm{d}p^4 \ .$$

In the new variables, the differential “volume” (which is different) is

$$\mathrm{d}v' = \mathrm{d}q^1\,\mathrm{d}q^2\,\mathrm{d}q^3\,\mathrm{d}q^4$$

and the equation that relates these two is the Jacobian,

$$\mathrm{d}v = J\mathrm{d}v' \ .$$

Thus, if we are working with the integral of a tensor T in the space of p^i's, $\int T\mathrm{d}v$ may be directly related to the space of g's if instead of T we use the tensor t, such that $\sqrt{|g|}\,t = T$. In fact,

$$\begin{aligned}\int T\mathrm{d}v &= \int \sqrt{|g_{rs}|}\;t\mathrm{d}v\\ &= \int \sqrt{|f_{ij}|}\;\frac{t}{J}\mathrm{d}v\\ &= \int \sqrt{|f_{ij}|}\;\frac{t'}{J}J\mathrm{d}v'\\ &= \int \sqrt{|f_{ij}|}\;t'\mathrm{d}v' \ .\end{aligned}$$

Obviously the tensor T has changed to T’, and so t to t’ after the change of variables.

0.2.6 Notes on the Jacobian

In changing n of the coordinates from $p \to q$, the variation of the point M changes from $\mathrm{d}p^1, \ldots, \mathrm{d}p^n$ to $\mathrm{d}q^1, \ldots, \mathrm{d}q^n$. Let us call F the transformation which relates both coordinate systems,

$$p^i = F^i(q^1, \ldots, q^n) \ .$$

In order to ensure the inverse relationship and its uniqueness, the Jacobian must be different from zero.

$$J = \begin{pmatrix} F^1_{q^1} & F^1_{q^2} & \ldots & F^1_{q^n}\\ F^2_{q^1} & F^2_{q^2} & \ldots & F^2_{q^n}\\ \vdots & \vdots & \ddots & \vdots\\ F^n_{q^1} & F^n_{q^2} & \ldots & F^n_{q^n}\end{pmatrix}$$

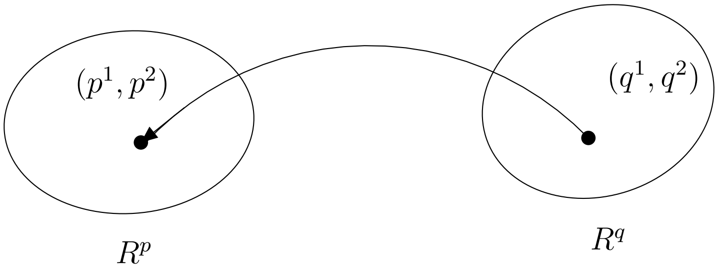
The partial derivative of F^i with respect to q^j is indicated by $F^i_{q^j}$. The differential change of the coordinates will be related by

$$\begin{pmatrix} \mathrm{d}p^1 \\ \vdots \\ \mathrm{d}p^n \end{pmatrix} = \begin{pmatrix} F^1_{q^1} & \dots & F^1_{q^n} \\ \vdots & \ddots & \vdots \\ F^n_{q^1} & \dots & F^n_{q^n} \end{pmatrix} \begin{pmatrix} \mathrm{d}q^1 \\ \vdots \\ \mathrm{d}q^n \end{pmatrix} .$$

Let us see this in two dimensional spaces

$$p^1 = F^1(q^1, q^2) \qquad \text{and} \qquad p^2 = F^2(q^1, q^2) \; . \tag{0.2.1}$$

The functions F^1 and F^2 create an affine transformation or a mapping which assigns every point (q^1, q^2) in the region R^q to a point (p^1, p^2) in the region R^p.



In matrix form this is

$$\begin{pmatrix} \mathrm{d}p^1 \\ \mathrm{d}p^2 \end{pmatrix} = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} \begin{pmatrix} \mathrm{d}q^1 \\ \mathrm{d}q^2 \end{pmatrix}$$

where the matrix (a_{ij}) expresses the linear transformation shown in (0.2.1) . Thus, R^p is the image of R^q through the matrix (a_{ij}). In order to assume that the inverse transformation exists and is unique, the determinant of the matrix (a_{ij}) should be non-zero.

Thus, in the affine transformation a line is transformed into a line and a surface into another surface.

In a differential form,

$$\begin{pmatrix} \mathrm{d}p^1 \\ \mathrm{d}p^2 \end{pmatrix} = \begin{pmatrix} F^1_{q^1} & F^1_{q^2} \\ F^2_{q^1} & F^2_{q^2} \end{pmatrix} \begin{pmatrix} \mathrm{d}q^1 \\ \mathrm{d}q^2 \end{pmatrix} = \begin{pmatrix} J \end{pmatrix} \begin{pmatrix} \mathrm{d}q^1 \\ \mathrm{d}q^2 \end{pmatrix}$$

where

$$F^i_{q^j} = \frac{\partial F^1(q^1, q^2)}{\partial q^j} \ .$$

Thus, the inverse transformation exists and is unique if

$$\begin{pmatrix} J \end{pmatrix} \neq 0 \ .$$

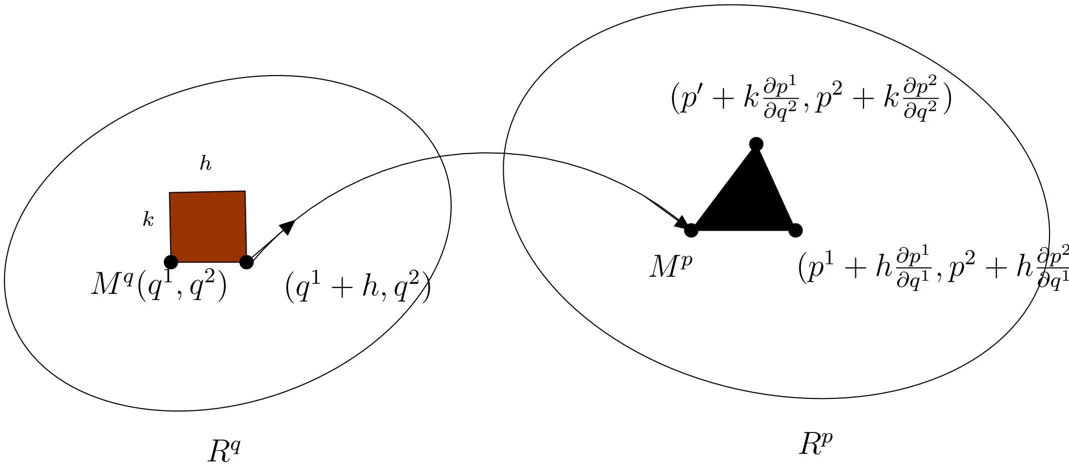
A classical application of the Jacobian is done in the calculation of multiple integrals when we change the variables in the integrand. Specifically, the integral defined in the space R^p is to be transformed to the region R^q. For simplicity, let us assume two dimensional regions,

$$I = \int_{R^p} F(p^1, p^2) \mathrm{d}p^1 \, \mathrm{d}p^2 \ .$$

Next we show that in the R^p region,

$$I = \int_{R^q} F(p^1(q^1, q^2), p^2(q^1, q^2)) \, J \, \mathrm{d}q^1 \, \mathrm{d}q^2$$

where J is the Jacobian of the coordinate transformation.


$(p' + k\frac{\partial p^1}{\partial q^2}, p^2 + k\frac{\partial p^2}{\partial q^2})$
h
k
$M^q(q^1, q^2)$
$(q^1 + h, q^2)$
M^p
$(p^1 + h\frac{\partial p^1}{\partial q^1}, p^2 + h\frac{\partial p^2}{\partial q^1}$
R^q
R^p

The element of area of R^p is (hk). The corresponding element of area in R^q is twice the shaded triangle. We recall that the area of the triangle of coordinates $(0,0), (a,b), (c,d)$ is $\frac{1}{2}(ad - bc)$. In the above figure the area of the shaded triangle is

$$\frac{1}{2}\left(\frac{\partial p^1}{\partial q^1}\frac{\partial p^2}{\partial q^2} - \frac{\partial p^1}{\partial q^2}\frac{\partial p^2}{\partial q^1}\right) hk \ .$$

The area of the parallelogram is twice this, so the ratio between both differential areas is just the quantity inside the parentheses, and this is the Jacobian of the transformation,

$$\mathrm{d}p^1\mathrm{d}p^2 = J\,\mathrm{d}q^1\mathrm{d}q^2 \ .$$

Differential Volume

As a simple generalization of the previous equations, the four-dimensional [redacted] volume is

$$\mathrm{d}v = \mathrm{d}p^1\mathrm{d}p^2\mathrm{d}p^3\mathrm{d}p^4 \ .$$

In the new variables, it is

$$\mathrm{d}v' = \mathrm{d}q^1\mathrm{d}q^2\mathrm{d}q^3\mathrm{d}q^4$$

and the equation that relates the two is the Jacobian,

$$\mathrm{d}v = J\mathrm{d}v' \ .$$

We have also seen that the metric tensor in both spaces is related by

$$\sqrt{|f_{rs}|} = J\sqrt{|g_{ij}|} \ .$$

Then, if we are working with the integral of a tensor T in the space of p's, $\int T\mathrm{d}v$ may be directly related to the space of q's by using the tensor t instead of T, such that $\sqrt{|g|}t = T$. In fact, $\int T\mathrm{d}v$ is

$$\int \sqrt{|g_{ij}|}t \ \mathrm{d}v = \int \sqrt{|f_{rs}|}\frac{t}{J} \ \mathrm{d}v = \int \sqrt{|f_{rs}|}\frac{t'}{J}J \ \mathrm{d}v' = \int \sqrt{|f_{rs}|}t' \ \mathrm{d}v' \ .$$

So, the tensor T has changed to T', and t to t' after the change of coordinates. d

If the p coordinates are chosen to be orthogonal, then their metric tensor g_{ij} has a determinant equal to 1. so, the metric tensor f_{rs} of the curvilinear coordinate q's has a determinant such that

$$\sqrt{|f_{rs}|} = J$$

where J is the Jacobian that expresses the transformation from p's to q's. Then, the element of volumen built by the hyper-parallepiped $\vec{e_1} \ \mathrm{d}q^1, \vec{e_2} \ \mathrm{d}q^2, \ldots$ is given by

$$\mathrm{d}v = \sqrt{|f_{rs}|} \ \mathrm{d}q^1 \ \mathrm{d}q^2 \ldots \ .$$

0.3 Lecture 3

0.3.1 Absolute Tensor Derivative

Most differential equations in physics contain first or second derivatives of some tensors. We must be sure that the derivative does not change the nature of the tensor. In other words, the derivative must also be a tensor. For instance, a vector, when moved differentially to a close point, must still be a vector. Thus if a vector $\vec{V}(q^1,\dots)$ is written as

$$\vec{V} = \vec{e}_i V^i$$

where $\vec{e}_i$ is the unit vector, and v^i its contravariant component, its total derivative will depend on the point where it exists. When it moves to the next point, both v^i *and* $\vec{e}_i$ may vary. Using the letter D for the "total" or "absolute differential",

$$\mathrm{D}\vec{V} = \mathrm{d}\vec{e}_i V^i + \mathrm{d}V^i \vec{e}_i$$

where $\vec{V}$ is a function of $q^i, q^2, \dots$ etc and also $\vec{e}_i$. The differential variation of $\vec{e}_i, \mathrm{d}\vec{e}_i$, may be written in a linear differential form

$$\mathrm{d}\vec{e}_i = \Gamma^j_{ki}\mathrm{d}q^k \vec{e}_j$$
$$\frac{\mathrm{d}\vec{e}_i}{\mathrm{d}q^k} = \Gamma^j_{ki}\vec{e}_j \ .$$

The connecting factor Γ^j_{ki} is called the Christoffel symbol.

Thus,

$$\mathrm{D}\vec{V} = \mathrm{d}V^i\vec{e}_i + V^i\Gamma^j_{ki}\mathrm{d}q^k\vec{e}_j \ .$$

Changing i for j in the first term,

$$\mathrm{D}\vec{V} = (\mathrm{d}V^j + V^i\Gamma^j_{ki}\mathrm{d}q^k)\vec{e}_j \ .$$

In general, a vector $\vec{W}$ is expressed as $\vec{W} = W^i\vec{e_{ij}}$, where W^i are its contravariant components. Thus, the contravariant component of $\mathrm{D}\vec{V}$ is the expression within the parentheses. We indicate it as

$$\mathrm{D}V^j = \mathrm{d}V^j + V^i\Gamma^j_{ki}\mathrm{d}q^k \ .$$

Since

$$\mathrm{d}V^j = \frac{\partial V^j}{\partial q^k}\mathrm{d}q^k \ ,$$

then

$$\mathrm{D}V^j = (\frac{\partial V^j}{\partial q^k} + V^i\Gamma^j_{ki})\mathrm{d}q^k \ .$$

So the absolute derivative of V expressed in its contravariant components is given by:

$$\mathrm{D}_k V^j = \frac{\partial V^j}{\partial q^k} + V^i\Gamma^j_{ki} \ .$$

As we will see next, the absolute derivative of $\vec{V}$ expressed by its contravariant components V_j along the coordinate q^k is:

$$\mathrm{D}_k V_j = \frac{\partial V_j}{\partial q^k} - V_h\Gamma^h_{ki} \ .$$

To prove the above, we assume that

$$\mathrm{D}_k V_j = \frac{\partial V_j}{\partial q^k} + B^h_{jk}V_h$$

where B^h_{jk} is to be determined.

Let us take a scalar made up from the contravariant component of $\vec{W}, W^i$, and the covariant component of $\vec{V}, V_i$. That is, the scalar S is defined as

$$S = W^i V_i \ .$$

Since S is a scalar, the absolute derivative is equal to its functional derivative. Thus,

$$\mathrm{D}_k S = \frac{\partial}{\partial q^k}(W^i V_i) = \frac{\partial W^i}{\partial q^k}V_i + W^i\frac{\partial V^i}{\partial q^k} \ . \qquad (0.3.1)$$

On the other hand, taking the absolute derivative of each component,

$$\mathrm{D}_k(W^iV_i) = (\mathrm{D}_kW^i)V_i + W^i(\mathrm{D}_kV_i) \ .$$

Introducing Γ and B in the last equation,

$$\mathrm{D}_k(W^iV_i) = V_i\partial_kW^i + V_i\Gamma^i_{kj}W^j + W^i\partial_kV_i + W^iB^j_{ki}V_j \ . \qquad (0.3.2)$$

Comparing (0.3.1) and (0.3.2), we see that the following sum must be 0:

$$V_i\Gamma^i_{kj}W^j + W^iB^j_{ki}V_j = 0 \ .$$

Changing the repeated indices, one finds that

$$W^iV_i(\Gamma^i_{kj} + B^j_{ki}) = 0 \ .$$

That is, $B^j_{ki} = -\Gamma^i_{kj}$.

Thus, the absolute derivative of the vector V in its covariant form is:

$$\mathrm{D}_kV_j = \partial_kV_j - \Gamma^i_{kj}V_i \ . \qquad (0.3.3)$$

Absolute Derivative of the Metric Tensor, g_{ij}

$$g_{ij} = \vec{e_i} \cdot \vec{e_j}$$

The variation of the unit vector $\vec{e_i}$ is given by the Christoffel symbol as:

$$\mathrm{d}\vec{e_i} = \Gamma^m_{ik}\vec{e_m}\delta q^k \ .$$

Thus, the variation of g_{ij} is:

$$\mathrm{d}g_{ij} = (\Gamma^l_{ki}\mathrm{d}q^k\vec{e_l}) \cdot \vec{e_j} + (\Gamma^m_{kj}\mathrm{d}q^k\vec{e_m}) \cdot \vec{e_i} \ ,$$

and since g_{ij} is a scalar,

$$\mathrm{d}g_{ij} = \frac{\partial g_{ij}}{\partial q^k}\mathrm{d}q^k \ .$$

From these last two equations we find:

$$\frac{\partial g_{ij}}{\partial q^k} = \Gamma^l_{ki} g_{jl} + \Gamma^m_{kj} g_{im} \ .$$

We could rewrite this equation by using the other form of Christoffel symbols:

$$\Gamma^l_{ki} g_{jl} = \Gamma_{kji} \quad \text{and} \quad \Gamma^m_{kj} g_{im} = \Gamma_{kij} \ .$$

Thus,

$$\frac{\partial g_{ij}}{\partial q^k} = \Gamma_{kji} + \Gamma_{kij} \ . \tag{0.3.4}$$

Christoffel Symbols as Functions of the Derivative of the Metric Tensor

Written

$$\frac{\mathrm{d}}{\mathrm{d}q^k} g_{ij} \quad \text{as} \quad g_{ij,k}$$

we rewrite three equations like (0.3.4) with different orders of the indices. i, k, j:

$$g_{ij,k} = \Gamma_{kij} + \Gamma_{ijk}$$

$$g_{ki,j} = \Gamma_{jki} + \Gamma_{kij}$$

$$g_{jk,i} = \Gamma_{ijk} + \Gamma_{jki}$$

Then,

$$g_{ij,k} + g_{ki,j} - g_{jk,i} = \Gamma_{kij} + \Gamma_{kij} = 2\Gamma_{jik}$$

Thus,

$$\Gamma_{jik} = \Gamma_{kij} = \frac{1}{2}(g_{ij,k} + g_{ki,j} - g_{jk,i})$$

We recall the symmetric properties of g and Γ, as well as their definitions:

$$g_{ij} = g_{ji}$$

$$d\vec{e_i} = \Gamma^l_{ki} \vec{e_l} dq^k$$

$$\Gamma_{ijk} = \Gamma_{kji}$$

$$\Gamma^k_{ij} g_{kl} = \Gamma_{ilj}$$

$$\Gamma_{ijk} g^{jl} = \Gamma^l_{ik}$$

Absolute Derivative of a Tensor of the Second Order (Two Indices), Tensor t_{ij}:

$$\mathrm{D}_k t_{ij} = t_{ij,k} = \frac{\partial t_{ij}}{\partial q^k} - \Gamma^m_{ki} t_{mj} - \Gamma^l_{kj} t_{il}$$

This equation may be derived by using the expression of the absolute derivative of a vector W_i such that $W_i = t_{ij}V^j$, where V^j is an arbitrary vector in its contravariant form and uniform in the space. By uniform, we mean

$$\frac{\partial V^i}{\partial q^k} = 0$$

and so its absolute derivative is given by

$$\mathrm{D}_k V^j = \Gamma^j_{kl} V^l$$

(The proof is shown in the next note.)

The equation above could be applied to the second order metric tensor g_{ij}. Thus,

$$\mathrm{D}_k g_{ij} = \frac{\partial g_{ij}}{\partial q^k} - \Gamma^n_{ki} g_{nj} - \Gamma^l_{kj} g_{il} = \frac{\partial g_{ij}}{\partial q^k} - \Gamma_{kji} - \Gamma_{kij} \ .$$

According to the equation derived earlier,

$$\frac{\partial g_{ij}}{\partial q^k} = \Gamma_{kji} + \Gamma_{kij}$$

and

$$\mathrm{D}_k g_{ij} = \frac{\partial g_{ij}}{\partial q^k} - \Gamma_{kji} - \Gamma_{kij} = \Gamma_{kji} + \Gamma_{kij} - \Gamma_{kji} - \Gamma_{kij} = 0 \ .$$

This conclusion is known as the Ricci Theorem or Ricci Identity.

Note on the Absolute Derivative of a Second Order Tensor t_{ij}:

Let us multiply it by the uniform vector V^j. Since uniform means that the vector has the same component in every point in space, i.e., $\dfrac{\partial V^j}{\partial q^k} = 0$, then

$$\mathrm{D}_k V^j = \frac{\partial V^j}{\partial q^k} + \Gamma^j_{ki} V^i = \Gamma^j_{ki} V^i \ .$$

Now the vector $t_{ij}V^j = W_i$ has an absolute derivative

$$\mathrm{D}_k W_i = \mathrm{D}_k t_{ij} V^j = \frac{\partial W_i}{\partial q^k} - \Gamma^l_{ik} W_l \ .$$

On the other hand, the absolute derivative of a product follows the same rules as the contravariant one

$$\mathrm{D}_k (t_{ij} V^j) = V^j \mathrm{D}_k t_{ij} + t_{ij} \mathrm{D}_k V^j = V^j \mathrm{D}_k t_{ij} + t_{ij} \Gamma^j_{km} V^m \ .$$

Equating these two previous expressions for $\mathrm{D}_k(t_{ij}V^j)$ we get

$$\frac{\partial W_i}{\partial q^k} - \Gamma^l_{ik} W_l = V^j \mathrm{D}_k t_{ij} + t_{ij} \Gamma^j_{km} V^m \ .$$

Since

$$\frac{\partial W_i}{\partial q^k} = \frac{\partial t_{ij} V^j}{\partial q^k} = \frac{\partial t_{ij}}{\partial q^k} V^j$$

and

$$W_l = T_{ls} V^s \ ,$$

we rewrite the previous equation as

$$\frac{\partial t_{ij}}{\partial q^k} V^j - \Gamma^l_{ik} t_{ls} V^s = V^j \mathrm{D}_k t_{ij} + t_{ij} \Gamma^j_{km} V^m \ .$$

Now changing the index s for j in the left hand side of the equation, and also j for m in the last term of the right hand side, we obtain

$$\mathrm{D}_k \left(t_{ij} \right) = \frac{\partial t_{ij}}{\partial q^k} - \left(\Gamma^m_{ik} t_{mj} - \Gamma^m_{kj} t_{im} \right) \ .$$

Summary of the Absolute Tensor Derivative of the Previous Examples

1. The scalar tensor is a simple function $F\left(q^1,\ldots,q^n\right)$. Its tensor derivative is just the common derivative that we learn in Calculus. Calling Q the tensor and $\mathrm{D}_k Q$ the tensor derivative with respect to the variable q^k:

$$Q = F \quad \text{and} \quad \mathrm{D}_k F = \frac{\partial \mathrm{F}\left(q^1,\ldots,q^n\right)}{\partial q^k}\ .$$

2. The next order of tensor is a vector $\vec{V}$. We say that the order is equal to the number of subindices. If we called this number m, then the scalar corresponds to $m = 0$. For this case, $m = 1$ and Q will have only one index. When it is up (down) it corresponds to the covariant (contravariant) components of the vector. In fact, we have seen that

$$\mathrm{D}_k V_i = \frac{\partial V_i}{\partial q^k} - \Gamma^j_{ik} V_j \quad \text{and} \quad \mathrm{D}_k V^i = \frac{\partial V^i}{\partial q^k} + \Gamma^i_{jk} V^j\ .$$

The element or tensor in question is $Q = \vec{V}$, $Q_i \equiv V_i$ and $Q^i \equiv V^i$.

3. Finally, a second order element Q was discussed, the tensor t_{ij} of two covariant indices. We found that

$$\mathrm{D}_k t_{ij} = \frac{\partial t_{ij}}{\partial q^k} - \Gamma^l_{kj} t_{il} - \Gamma^m_{ik} t_{mj}\ .$$

In general if Q is a covariant tensor of order m, that is, if Q is expressed as $Q_{r_1,\ldots,r_m}$, then its absolute tensor derivative is

$$\mathrm{D}_k Q_{r_1,\ldots,r_m} = \frac{\partial Q_{r_1,\ldots,r_m}}{\partial q^k} - \sum_{r_l=r_1}^{r_l=r_m} \Gamma^s_{r_l k} Q_{r_1,\ldots,r_{l-1},s,r_{l+1},\ldots,r_m}\ .$$

For the contravariant form of Q,

$$\mathrm{D}_k Q^{r_1,\ldots,r_m} = \frac{\partial Q_{r_1,\ldots,r_m}}{\partial q^k} + \sum_{r_l=r_1}^{r_l=r_m} \Gamma^{r_l}_{sk} Q^{r_1,\ldots,r_{l-1},s,r_{l+1},\ldots,r_m} .$$

The summation in both equations with respect to s from 1 to n is understood and there are m terms with Christoffel symbols.

These equations for the absolute derivative were introduced by M. Christoffel in 1869.

4. Relation between Γ and the second derivative of the variation of the point M in the space.

$$\mathrm{d}\overrightarrow{M} = \overrightarrow{e}_i \mathrm{d}q^i$$

So $\mathrm{d}q^i$ are the contravariant components of $\mathrm{d}\overrightarrow{M}$. Similarly

$$\mathrm{d}\overrightarrow{e}_i = \left(\mathrm{d}q^k \Gamma^l_{ki}\right) \overrightarrow{e}_l$$

so $\left(\mathrm{d}q^k \Gamma^l_{ki}\right)$ are the contravariant components of $\mathrm{d}\overrightarrow{e}_i$. From those equations

$$\frac{\mathrm{d}\vec{M}}{\mathrm{d}q^i} = \frac{\partial^2 \overrightarrow{M}}{\partial q^i \partial q^k} \vec{e}_i$$

and so the Christoffel symbol is directly related to the second derivative of the point $\vec{M}$ as it moves in the space.

5. The absolute derivative implies the variation of the unit vector $\overrightarrow{e}_i$ and it is indicated by D. The small d in $\mathrm{d}V^i$ is the partial

derivative of the component V^i of $\overrightarrow{V}$ multiplied by the variation of the coordinate

$$\mathrm{d}V^i = \frac{\partial V^i}{\partial q^k}\mathrm{d}q^k$$

NOTE: Advanced Conclusions on Parallel Transport as it will be Discussed in the Next Section

We may see that the Ricci relationship, $\mathrm{D}_k g_{ij} = 0$, implies that the length of a vector that moves parallel to itself does not change. In fact, the length l of a vector $\overrightarrow{V}$ is given by

$$l = V^i V^j g_{ij}$$

so

$$\mathrm{d}l = \left(DV^i\right) V^j g_{ij} + V^i \left(DV^j\right) g_{ij} + V^i V^j D g_{ij}$$

We indicate the tensor or absolute derivative by D. In the case of l, which is a scalar, D is the same as d, which indicates the simple derivative of a scalar function.

Now according to what will be discussed in the next section, the absolute derivative of $\overrightarrow{V}$ as it is displaced parallel to itself is null. Thus, the two first terms are zero and

$$\mathrm{d}l = V^i V^j D g_{ij}$$

Thus, if $Dg_{ij} = 0$ then $\mathrm{d}l = 0$ as well.

The Ricci relationship applies to Riemann space and of course to Euclidean space, but for the Weyl space there is a more general relationship. In fact, Weyl space allows for a variation of length as $\overrightarrow{V}$ moves parallel to itself. The ”connecting function”, or equivalently, the new Christoffel symbols, are different from the one discussed previously.

In fact, we will see that in any displacement

$$D_m g_{ij} = \frac{\partial g_{ij}}{\partial q^m} + \Gamma_{mij} + \Gamma_{imj} = g_{ij}\Phi_m \ .$$

The last equation express the linearization of dl, which according to Weyl, is assumed when the vector moves parallel to itself

$$\frac{\mathrm{d}l}{l} = \Phi_m \mathrm{d}q^m \tag{0.3.5}$$

The term $\bar{\Gamma}$ is equal to the Christofel symbol (Γ) plus a term that includes Φ_m. The relationship, as we will see, is

$$\bar{\Gamma}^k_{ij} = \Gamma^k_{ij} - \frac{1}{2}\left(\Phi_i g^k_j + \Phi_j g^k_i - \Phi_k g_{ij}\right) \ .$$

We will prove this in the next section.

We see that $D_m g_{ij} = 0$ means that $\Phi_m = 0$ and so d$l = 0$. The vector moves without changing its length. This conclusion assumes the definition of Φ_m, see (0.3.5). Thus, in the Weyl space, the variation of length of a vector which moves parallel to itself may be non-zero, and the following equation is valid:

$$D_m g_{ij} - g_{ij}\Phi_m = 0 \ .$$

This may be seen as a generalization of the Ricci Theorem.

Summary: Relationship Between the Christoffel Symbols and the Metric Tensor

$$\Gamma_{kji} = \frac{1}{2}(g_{kj,i} + g_{ji,k} - g_{ki,j})$$

$$\Gamma_{kji} = \Gamma_{ijk}$$

$$\Gamma^j_{ki} = g^{jh}\Gamma_{khi}$$

$$g^{jh} = (g^{-1})_{jh} = \frac{\text{cofactor}(g_{jh})}{D}$$

$$D = \text{Determinant of } g_{jh}$$

$$g_{jh} = g_{hj} \quad , \quad g_{lm,n} = \frac{\partial g_{lm}}{\partial q^n}$$

Examples of Christoffel Values for Curvilinear Coordinates

1. Polar Coordinates (ρ, θ)

   As it was before.

$$g_{\rho\theta} = \begin{pmatrix} 1 & 0 \\ 0 & \rho^2 \end{pmatrix}$$

and

$$g^{\rho\theta} = \begin{pmatrix} 1 & 0 \\ 0 & \frac{1}{\rho^2} \end{pmatrix}$$

so the square of the path of an element in the space is

$$\mathrm{d}s^2 = \mathrm{d}\rho^2 + \rho^2\mathrm{d}\theta^2$$

We will calculate Γ_{ijk} (i, j, k being ρ or θ). That is, we will find the Christoffel values for $\Gamma_{\rho\rho\rho}, \Gamma_{\rho\theta\rho}, \Gamma_{\rho\theta\theta}, \Gamma_{\theta\theta\rho}, \Gamma_{\theta\theta\theta}$. We know that Γ is symmetric in the extreme indices: $\Gamma_{\rho\theta\theta} = \Gamma_{\theta\theta\rho}$. Thus,

$$\Gamma_{\rho\rho\rho} = \frac{1}{2}(g_{\rho\rho,\rho} + g_{\rho\rho,\rho} - g_{\rho\rho,\rho}) = \frac{1}{2}g_{\rho\rho,\rho} = 0$$

$$\Gamma_{\rho\theta\rho} = \frac{1}{2}(g_{\rho\theta,\rho} + g_{\theta\rho,\rho} - g_{\rho\rho,\theta}) = 0$$

$$\Gamma_{\rho\theta\theta} = \frac{1}{2}(g_{\rho\theta,\theta} + g_{\theta\theta,\rho} - g_{\rho\theta,\theta}) = \frac{1}{2}\frac{\partial\rho^2}{\partial\rho} = \rho$$

$$\Gamma_{\theta\theta\rho} = \frac{1}{2}(g_{\theta\theta,\rho} + g_{\rho\theta,\theta} - g_{\theta\rho,\theta}) = \frac{1}{2}\frac{\partial\rho^2}{\partial\rho} = \rho$$

$$\Gamma_{\theta\theta\theta} = \frac{1}{2}(g_{\theta\theta,\theta} + g_{\theta\theta,\theta} - g_{\theta\theta,\theta}) = \frac{1}{2}g_{\theta\theta,\theta} = \frac{1}{2}\frac{\partial\rho^2}{\partial\theta} = 0$$

$$\Gamma_{\theta\rho\theta} = \frac{1}{2}(g_{\theta\rho,\theta} + g_{\rho\theta,\theta} - g_{\theta\theta,\rho}) = -\frac{1}{2}\frac{\partial\rho^2}{\partial\rho} = -\rho$$

$$\Gamma_{\theta\rho\rho} = \frac{1}{2}(g_{\theta\rho,\rho} + g_{\rho\rho,\theta} - g_{\rho\theta,\rho}) = 0$$

2. Spherical Coordinates

$$g_{\rho\theta\psi} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \rho^2 & 0 \\ 0 & 0 & \rho^2\sin^2\theta \end{pmatrix} \quad \text{where } i,j = \rho,\theta,\psi$$

$$\mathrm{d}s^2 = \mathrm{d}\rho^2 + \rho^2\mathrm{d}\theta^2 + \sin^2\theta\mathrm{d}\psi^2$$

$$\Gamma_{\rho\rho\rho} = \frac{1}{2}(g_{\rho\rho,\rho} + g_{\rho\rho,\rho} - g_{\rho\rho,\rho}) = \frac{1}{2}g_{\rho\rho,\rho} = 0$$

$$\Gamma_{\theta\theta\rho} = \frac{1}{2}(g_{\theta\theta,\rho} + g_{\theta\rho,\theta} - g_{\rho\theta,\theta}) = \frac{1}{2}\frac{\partial\rho^2}{\partial\rho} = \rho$$

$$\Gamma_{\psi\psi\theta} = \rho^2\sin\theta\cos\theta$$

$$\Gamma_{\psi\rho\rho} = 0$$

$$\Gamma_{\rho\psi\psi} = \rho\sin^2\theta$$

$$\Gamma_{\rho\theta\theta} = \rho$$

Christoffel Symbols in Polar Coordinates and Their Relationship to the Centripetal Force

From the values of Γ_{ijk}, where i,j,k are ρ or θ, we determine Γ^i_{kj} such that

$$\Gamma_{kjl} = \Gamma^i_{kj}g_{il} \quad \text{or} \quad \Gamma^i_{kl} = g^{ji}\Gamma_{kjl}\ .$$

We recall that $g^{\rho\rho} = 1$, $g^{\theta\theta} = \dfrac{1}{\rho^2}$ and that otherwise $g^{ij} = 0$. Let us determine the following: $\Gamma^\rho_{\rho\rho}$, $\Gamma^\rho_{\theta\rho}$, $\Gamma^\rho_{\theta\theta}$

$$\Gamma^{\rho}_{\rho\rho} = g^{\rho\rho}\Gamma_{\rho\rho\rho} + g^{\rho\theta}\Gamma_{\rho\theta\rho} = 0$$

$$\Gamma^{\rho}_{\theta\rho} = g^{\rho\rho}\Gamma_{\theta\rho\rho} + g^{\rho\theta}\Gamma_{\theta\theta\rho} = 0$$

$$\Gamma^{\rho}_{\theta\theta} = g^{\rho\theta}\Gamma_{\theta\theta\theta} + g^{\rho\rho}\Gamma_{\theta\rho\theta} = -\rho$$

Calling γ the acceleration of a particle around a curve whose parameter is t, its contravariant components are given by:

$$\gamma^i = \frac{\mathrm{d}\zeta^i}{\mathrm{d}t} + \Gamma^i_{kj}\zeta^k \frac{\mathrm{d}q^j}{\mathrm{d}t}$$

where $\zeta^i = \dfrac{\mathrm{d}q^i}{\mathrm{d}t}$ and i is θ or ρ. So $\zeta^\theta = \dot{\theta}$ and $\zeta^\rho = \dot{\rho}$. Assuming $i = \rho$, then

$$\gamma^\rho = \ddot{\rho} + \Gamma^{\rho}_{kj}\zeta^k \dot{q}^j = \ddot{\rho} + \Gamma^{\rho}_{\theta\theta}\zeta^\theta\dot{\theta}$$

Where $\Gamma^{\rho}_{\theta\theta} = 0$ if and only if $\Gamma^{\rho}_{XY} = 0$. Thus,

$$\gamma^\rho = \ddot{\rho} - \rho\dot{\theta}^2 = \ddot{\rho} - \frac{v^2}{\rho}$$

Where $v = \rho\dot{\theta}$. The last term corresponds to the centripetal acceleration.

0.4 Lecture 4

0.4.1 Riemann Space

Let us look into the absolute derivative in relation to parallel displacement in the Riemann Space.

Let us take one point M in the space that will be a function of $q^1, q^2, q^3, \ldots$ For simplicity, we use only q to represent $q_1, q_2, \ldots$ A vector applied at that point may keep the same component as it moves parallel to itself. That is,

$$V^i(q+\delta q) = V^i(q) \ .$$

Although this may be true for a given system of coordinates, it is no necessarily so for all coordinate systems. Let us assume that it is correct for the system $\bar{q}$. This system is called a system of geodesic coordinates. Thus,

$$V^i(\bar{q}+\delta\bar{q}) = V^i(\bar{q}) \ .$$

Now, in a coordinate transformation from the system $\bar{q}$ to q,

$$V^i(q+\delta q) = A^i_k(\bar{q}+\delta\bar{q}) V^k(\bar{q}+\delta\bar{q})$$

where A^i_k is the matrix transformation equal to

$$\frac{\partial q^i}{\partial \bar{q}^k} \ .$$

Neglecting second terms we obtain:

$$V^i(q+\delta q) = \Big(A^i_k(\bar{q}) + \frac{\partial A^i_k}{\partial \bar{q}^h}\delta\bar{q}^h\Big)V^k(\bar{q}) \ .$$

In the last term we used the previous equality $V^k(\bar{q}+\delta\bar{q}) = V^k(\bar{q})$. Now, since

$$V^i(q) = V^k(\bar{q})A^i_k(\bar{q})$$

then

$$V^i(q+\delta q) - V^i(q) = \frac{\partial A^i_k}{\partial \bar{q}^h} V^k(\bar{q})\delta\bar{q}^h \ .$$

Thus, the vector $\vec{V}$ that moved parallel to itself has equal components in the geodesic, but different ones in another curve given by the previous equation. So the vector $\vec{V}$ moves parallel to itself as long as the difference of its components is equal to (replacing A^i_k by $\dfrac{\partial q^i}{\partial \bar{q}^k}$)

$$\frac{\partial^2 q^i}{\partial \bar{q}^k \partial \bar{q}^h} V^k(\bar{q}) \delta \bar{q}^h \ .$$

Now let us express the difference in the q system of coordinates using the inverse transformation

$$A'^l_m(q) = \frac{\partial \bar{q}^l}{\partial q^m} \ .$$

So,

$$V^k(\bar{q}) = A'^k_m(q) V^m(q)$$

and

$$\delta \bar{q}^h = A'^h_l(q) \delta q^l \ .$$

Then,

$$V^i(q + \delta q) - V^i(q) = \left[\frac{\partial^2 q^i}{\partial \bar{q}^k \partial \bar{q}^h} A'^k_m A'^h_l \right] V^m \delta q^l \ .$$

By definition,

$$\Gamma^i_{ml} = -\frac{\partial^2 q^i}{\partial \bar{q}^k \partial \bar{q}^h} A'^k_m A'^h_l \ .$$

Γ^i_{ml} is symmetric with respect to m and l. if the space admits curves with torsion, this not true anymore and we are in the space of Cartan.

In summary, a parallel displacement of the contravariant components of the vector are not necessarily preserved, but they change as

$$V^i(q + \delta q) = V^i(q) - \Gamma^i_{ml} V^m \delta q^l \ .$$

That is, the variation of the components between two parallel vectors is not zero anymore, but rather is equal to

$$\frac{\partial V^i}{\partial q^l} \delta q^l = -\Gamma^i_{ml} V^m \delta q^l \ .$$

In other words, the total differential $\mathrm{D}V^i$ is

$$\mathrm{D}V^i(q) = \frac{\partial V^i}{\partial q^l}\delta q^l + \Gamma^i_{ml}V^m \delta q^l = 0 \ .$$

This result agrees with the previous section, where the Christoffel symbols were introduced to account for the variation of the unit vector $\vec{e_i}$. It tells us that the absolute derivative $\mathrm{D}V$ of the vector $\vec{V}$ is zero.

Let us introduce unit vectors again:

$$\vec{V} = \vec{e_i}V^i$$

The total or absolute derivative, $\mathrm{D}V$ (its component, $\dfrac{\mathrm{D}V}{\partial q^k}$, is called D_k), is:

$$\mathrm{D}_k\vec{V} = \frac{\partial V^i}{\partial q^k}\vec{e_i} + V^i\mathrm{D}_k\vec{e_i}$$

Since, in parallel transport, $\mathrm{D}_k\vec{V} = 0$, and we define Γ^m_{ik} as equal to the variation of the unit vector between two near points, $\mathrm{D}_k\vec{e_i} = \Gamma^m_{ik}\vec{e_m}$, we again find that

$$\frac{\partial V^i}{\partial q^k}\vec{e_i} + V^i\Gamma^m_{ik}\vec{e_m} = 0$$

Now, changing i for m in the first term, we get,

$$\frac{\partial V^m}{\partial q^k} + V^i\Gamma^m_{ik} = 0$$

which is the form for the derivative of the contravariant components of the vector $\vec{V}$ equated to zero, that we previously found by transporting $\vec{V}$ parallel to itself.

We summarize out previous pages by saying:

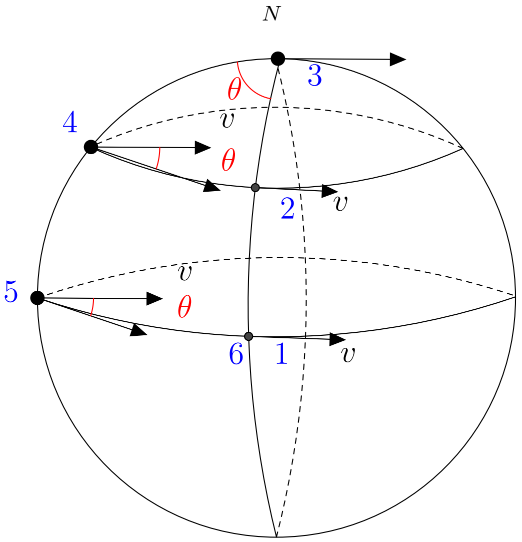
1. Parallel displacement or transport of a vector means that its absolute derivative (tensor derivative) is equal to zero.

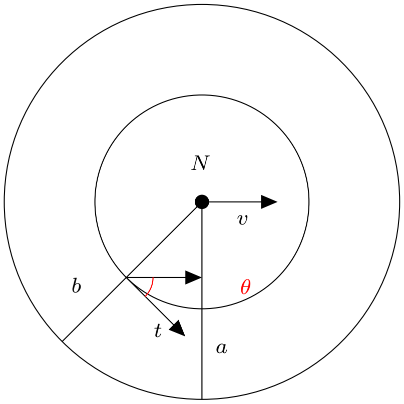
2. Parallel transport does not mean that the component of the vector is invariant except when it moves on a geodesic.

3. The variation of $\vec{e_i}$ from a point p (of coordinates $q^1, q^2, \ldots, q^n$) to a close point of coordinates $q^1 + \delta q^1, \ldots, q^n + \delta q^n$, is given by the expression $\mathrm{d}\vec{e_i} = \Gamma^m_{ik}\vec{e_m}$ as q^k varies. We found that as $\vec{V} = V^i\vec{e_i}$ moves parallel to itself, it obeys an equation which is consistent with this expression for $\mathrm{d}\vec{e_i}$. Thus, the Christoffel symbols are related to the second derivative of the variable qs with respect to the geodesic coordinates.

Example of a Geodesic and Parallel Transport

Assume a sphere and let us move a vector $\vec{V}$ on the surface if such sphere:





Let us start at 1 with $\vec{V}$ tangent to the equator, which is a geodesic. Let us move $\vec{V}$ to the North Pole perpendicularly to the meridian a. Thus, $\vec{V}$ is always perpendicular to a and will now form an angle with the tangent to any parallel on the b meridian. Therefore, when we reach the equation at 6, $\vec{V}$ is no longer tangent to the equator, but rather forms an angle θ with the initial position 1. So, in parallel transport from 1 to 6, $\vec{V}$ at the 6 position does not coincide with itself at the the 1 position.

In a plane or flat space, things are different. If we transport $\vec{V}$ on a circle with a constant angle θ with respect to the tangent, we always come back to the same initial vector. The difference is based on the fact that this space is flat while in the precious case it was curved. In a flat space, we may use curvilinear coordinates that we may transform into rectilinear coordinates where $\vec{V}$ may run parallel to itself with changing its components. They are the geodesic coordinates mentioned above.

Curvature of the Space and Riemann Space

As we discussed previously in the first section of 0.3.1, the variation of the unit vector $\mathrm{d}\vec{e_i}$ may be written as:

$$\mathrm{d}\vec{e_i} = (\Gamma^j_{ik}\mathrm{d}q^k)\vec{e_j} \tag{0.4.1}$$

where Γ^{j}_{ik} are the Christoffel symbols. This was discussed in the framework of a Euclidean space. The Riemann Space is more general. In this space, there is also a metric tensor g_{ij} and so a differential quadratic $\mathrm{d}s^2$ which is just a Pythagorean form between two close points:

$$\mathrm{d}s^2 = g_{ij}\mathrm{d}q^i\mathrm{d}q^j$$

where $(q^1, q^2, \ldots, q^n)$ are the contravariant coordinates.

The coefficients of g_{ij} are now any arbitrary functions of q^i, constrained to the fact that they are continuous differentiable functions up to a high order. There may not exist, however, any curvilinear coordinates that will generate the functions g_{ij}.

We may still work as we have done so far, assuming a Euclidean space at least locally. That is, we assume that it is always possible to have a Euclidean plane at the point $M(q^1, \ldots, q^n)$ of the Riemann Space, which is "tangent" to it at M and where the Euclidean metric g_{ij} at that point has the same value as that of $g_{ij}(M)$ of the Riemann Space.

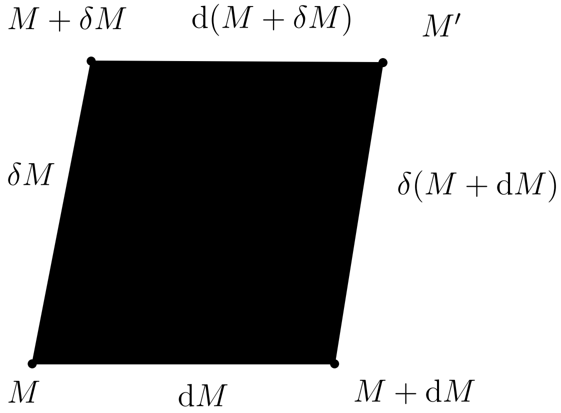
Every tensor at M has a corresponding one in the Euclidean "tangent" Space. The tensor algebra, which implies derivatives, is developed in this Euclidean space which is taken as representative of a local surroundings of the Riemann Space. At the point M, the algebra which characterizes the local geometry is the same in the Riemann Space as well as in the Euclidean "tnagent" plane. We say that the metrics are osculatory. The osculatrix plane will contain the normal vector to the tangent plane. A curve C passing through the point M in the Riemann Space will have been reproduced locally in the osculatrix plane with the same curvature that C has at M in the Riemann Space.

When moving to another point M', another Euclidean "tangent" space is considered with a different $\bar{g}'_{ij}$, such that

$$g_{ij}(M) = \bar{g}'_{ij} \ .$$

The algebraic connection between a point M and another $M' = M + \mathrm{d}M$, which carried out the absolute derivative, is done as before by the Christoffel symbols of the Euclidean Space tangent at M.

The curvature of the Riemann Space at each point may be calculated by considering an elemental parallelogram based on $M, M + \mathrm{d}M$, and $M + \delta M$:



We displace the unit vector at $M, \vec{e_i}$, from M to M' via two paths. Path 1 goes from M to $M + \delta M$ to M' and Path 2 goes from M to $M + \mathrm{d}M$ and then to M'. If the space is curved, the variations of $\vec{e_i}$ will be different on each path. If the difference is zero, then the space is flat.

The differences are taken in the Euclidean "tangent plane." Calling $\vec{e_i}'$ the "unit vector" at M' and $\vec{e_i}$ at M, then:

$$(\vec{e_i}' - \vec{e_i})_2 = \mathrm{d}\vec{e_i} + \delta\vec{e_i} + \delta\mathrm{d}\vec{e_i}$$

$$(\vec{e_i}' - \vec{e_i})_1 = \delta\vec{e_i} + \mathrm{d}\vec{e_i} + \mathrm{d}\delta\vec{e_i}$$

where the subscripts 1 and 2 stand for path 1 and path 2, respectively. Thus,

$$(\vec{e_i}')_2 - (\vec{e_i})_1 = \mathrm{d}\delta\vec{e_i} - \delta\mathrm{d}\vec{e_i}$$

When these two elementary vector variations d and δ commute, we have a perfect "plane"; otherwise these differences give rise to a tensor as follows. Since

$$\mathrm{d}\vec{e_i} = \Gamma^j_{ik}\mathrm{d}q^k\vec{e_j} \quad \text{and} \quad \delta\vec{e_i} = \Gamma^j_{ik}\delta q^k\vec{e_j}$$

then

$$\delta \mathrm{d}\vec{e_i} = \Gamma^j_{ik}\delta\mathrm{d}(q^k)\vec{e_j} + \Gamma^j_{ik}\mathrm{d}q^k\delta\vec{e_j} + (\partial_r\Gamma^j_{ik})\delta q^r\mathrm{d}q^k\vec{e_j}\ .$$

Similarly for $\mathrm{d}\delta\vec{e_i}$ (we change r for s)

$$\mathrm{d}\delta\vec{e_i} = \Gamma^j_{ik}\mathrm{d}\delta(q^k)\vec{e_j} + \Gamma^j_{ik}\delta q^k\mathrm{d}\vec{e_j} + (\partial_s\Gamma^j_{ik})\delta q^k\mathrm{d}q^s\vec{e_j}$$

In the difference of

$$\mathrm{d}\delta\vec{e_i} - \delta\mathrm{d}(\vec{e_i}) = (\mathrm{d}\delta - \delta\mathrm{d})\vec{e_i}$$

the first terms cancel out and we have

$$(\mathrm{d}\delta-\delta\mathrm{d})\vec{e_i} = (\Gamma^j_{ik}\delta q^k\mathrm{d}\vec{e_i}-\Gamma^j_{ik}\mathrm{d}q^k\delta\vec{e_i})+(\partial_s(\Gamma^j_{ik})\delta q^k\mathrm{d}q^s-\partial_r(\Gamma^j_{ik})\delta q^r\mathrm{d}q^k)\vec{e_j}\ .$$ [redacted]

In the last term of the last parentheses, we change k for s and r for k, so it becomes:

$$(\partial_s\Gamma^j_{ik} - \partial_k\Gamma^j_{is})\delta q^k\mathrm{d}q^s\vec{e_j}\ ,$$

while the first parentheses becomes, after introducing

$$\delta\vec{e_j} = \Gamma^m_{jl}\delta q^l\vec{e_m} \quad \text{and} \quad \mathrm{d}\vec{e_j} = \Gamma^m_{jl}\mathrm{d}q^l\vec{e_m}\ ,$$

$$(\Gamma^j_{ik}\Gamma^m_{jl} - \Gamma^j_{il}\Gamma^m_{jk})\mathrm{d}q^l\delta q^k\vec{e_m}$$

In the second term inside the parentheses we changed k for l and vice versa. If in the previous expression, which contains the derivative of the Christoffel symbols, we change j for m and s for l, we find that

$$(\mathrm{d}\delta - \delta\mathrm{d})\vec{e_i} = [(\partial_l\Gamma^m_{ik} - \partial_k\Gamma^m_{il}) + (\Gamma^j_{ik}\Gamma^m_{jl} - \Gamma^j_{il}\Gamma^m_{jk})]\vec{e_m}\mathrm{d}q^l\delta q^k = \Omega^m_i\vec{e_m}$$

where

$$\Omega^m_i = [(\partial_l\Gamma^m_{ik} - \partial_k\Gamma^m_{il}) + (\Gamma^j_{ik}\Gamma^m_{jl} - \Gamma^j_{il}\Gamma^m_{jk})]\mathrm{d}q^l\delta q^k$$

The curvature tensor is the expression in brackets:

$$R^m_{i,kl} = [(\partial_l\Gamma^m_{ik} - \partial_k\Gamma^m_{il}) + (\Gamma^j_{ik}\Gamma^m_{jl} - \Gamma^j_{il}\Gamma^m_{jk})]$$

For $R^m_{i,kl} = 0$, the operators d and δ commute, such that $(\mathrm{d}\delta - \delta\mathrm{d})\vec{e_i} = 0$.

By making equal the indices m and k we reduce the four -index tensor to a two-index one. The result is the tensor R_{il}, called the Ricci Tensor.

It is not difficult to show directly, using the coordinate transformation, that Ω^j_i is a mixed tensor and that $R^m_{i,kl}$ is as well. By simply recalling the definition of a tensor with regards to the coordinate transformation, you will find that both elements fulfill those conditions.

The tensor $R^m_{i,lk}$ is called the Riemann-Christoffel Tensor, and may be written with four covariant indexes as $R_{hi,lk}$ by multiplying $R^m_{i,lk}$ by g_{mh}. Let us rewrite the Riemann-Christoffel Tensor in the following way:

$$R^m_{i,lk} = (\partial_l \Gamma^m_{ik} + \Gamma^j_{ik}\Gamma^m_{jl}) - (l \leftrightarrow k)$$

where the last parentheses signify the first expression in parentheses with the indexes l and k interchanged. Thus,

$$R_{hi,lk} = g_{mh}[(\partial_l \Gamma^m_{ik} + \Gamma^j_{ik}\Gamma^m_{jl}) - (l \leftrightarrow k)]$$

We recall that

$$\partial_l g_{mh} = \Gamma_{lhm} + \Gamma_{lmh}$$

Therefore,

$$\partial_l(\Gamma^m_{ik} g_{mh}) = \partial_l(\Gamma_{ihk}) = \partial_l(\Gamma^m_{ik}) g_{mh} + \Gamma^m_{ik}\partial_l g_{mh}$$

So,

$$R_{hi,lk} = [\partial_l(\Gamma_{ihk}) - \Gamma^m_{ik}(\Gamma_{lhm} + \Gamma_{lmh}) + \Gamma^j_{ik}\Gamma_{jhl}] - [l \leftrightarrow k]$$

which simplifies to

$$R_{hi,lk} = [\partial_l \Gamma_{ihk} - \Gamma^m_{ik}\Gamma_{lmh}] - [l \leftrightarrow k]$$

where

$$\Gamma_{ihk} = \Gamma^{m}_{ik} g_{hm}, \text{etc.}$$

A Γ with three sub-indexes ($\Gamma_{...}$) is called a Christoffel symbol of the first kind, while a Γ with one super-index and two sub-indexes ($\Gamma^{\cdot}_{\cdot\cdot}$) is called a Christoffel symbol of the second kind.

Space With Normal Coordinates

As was mentioned when the concept of the geodesic was discussed, it is always possible to generate in the Euclidean Space a local rectilinear system of coordinates tangent at M called *normal coordinates*. In that system all metric tensor components g_{ij} are constants, and therefore the Christoffel symbols at that point are zero. In this system the curvature tensor $R_{hi,lk}$ is reduced to

$$R_{hi,lk} = \frac{1}{2}(\partial_l \partial_i g_{hk} + \partial_k \partial_h g_{il} - \partial_l \partial_h g_{ik} - \partial_k \partial_i g_{hl}) \ . \qquad (0.4.2)$$

In this equation, whose origin is the variation from point to point of the Christoffel symbols, we have used the expression for those Christoffel symbols as a function of the metric tensor.

The symmetric relationships can be easily seen:

$$R_{hi,lk} = R_{lk,hi}$$

$$R_{hi,lk} = -R_{hi,kl} = -R_{ih,lk}$$

and by circular permutation of the indexes i, l, k, it results that

$$R_{hi,lk} + R_{hl,ki} + R_{hk,il} = 0 \ . \qquad (0.4.3)$$

Ricci Tensor

At this point, it will be illuminating to review tensor contraction that reduces a tensor to that of two indexes. For any mixed tensor, equating a contravariant index to a covariant one means that we are contracting those indexes or reducing the tensor to two orders. For example,

$$T^{jk}_{il} \to T^{jk}_{ik} = (\sum_{k=1}^{n} T^{jk}_{ik}) = T^{j}_{i} \tag{0.4.4}$$

The parentheses emphasize that repeated indexes imply summation from 1 to n, where n is the number of dimensions in the Riemann Space. In relativity, $n = 4$.

The Ricci Tensor R_{rl}, whose source is the curvature tensor $R^{j}_{r,ml}$, is defined as

$$R_{rl} = \sum_{j} R^{j}_{r,jl} \tag{0.4.5}$$

R_{rl} is equal to R_{lr}, because

$$R_{rl} = g^{m}_{j} R^{j}_{r,ml} = g^{m}_{j} g^{jh} R_{hr,ml} = g^{m}_{j} g^{jh} R_{lm,rh} = g^{m}_{j} R^{j}_{lm,r} = R_{lr} \tag{0.4.6}$$

In the third step, we exchanged $ml \to hk$, $ml \to lm$, and $hr \to rh$, and then used the symmetric properties of $R_{\alpha\beta,\gamma\delta}$.

The Ricci Tensor may be reduced further to a simple scalar R by multiplying it by g^{hi}:

$$g^{hi} R_{hi} = R$$

R is called the scalar Riemann curvature.

Bianchi Identity

The equation with which we previously defined the curvature tensor $R^{m}_{i,lk}$ could also be written as

$$(\mathrm{D}_l \mathrm{D}_k - \mathrm{D}_k \mathrm{D}_l)\vec{e_i} = R^{m}_{i,kl} \vec{e_m} \tag{0.4.7}$$

where D_l indicates the absolute tensor or covariant derivative with respect to q^l. Now introducing the operator D_r in (0.4.7), we have

$$(\mathrm{D}_l \mathrm{D}_k \mathrm{D}_r - \mathrm{D}_k \mathrm{D}_l \mathrm{D}_r)\vec{e_i} = \mathrm{D}_r(R^{m}_{i,lk})\vec{e_m} + R^{m}_{i,lk}(\mathrm{D}_r \vec{e_m}) \tag{0.4.8}$$

If $t_{ij} = \vec{e_i}\vec{e_j}$, the external product of the vectors $\vec{e_i}$ and $\vec{e_j}$, then

$$\mathrm{D}_l \mathrm{D}_k (t_{ij}) = (\mathrm{D}_l \mathrm{D}_k \vec{e_i})\vec{e_j} + \vec{e_i}(\mathrm{D}_l \mathrm{D}_k \vec{e_j}) + (\mathrm{D}_k \vec{e_i})(\mathrm{D}_l \vec{e_j}) + (\mathrm{D}_l \vec{e_i})(\mathrm{D}_k \vec{e_j})$$

Changing $l \leftrightarrow k$, and the last two terms cancel out by subtraction. Thus,

$$(\mathrm{D}_l\mathrm{D}_k - \mathrm{D}_k\mathrm{D}_l)t_{ij} = \vec{e_j}(\mathrm{D}_l\mathrm{D}_k - \mathrm{D}_k\mathrm{D}_l)\vec{e_i} + \vec{e_i}(\mathrm{D}_l\mathrm{D}_k - \mathrm{D}_k\mathrm{D}_l)\vec{e_j}$$

Now, using (0.4.7),

$$(\mathrm{D}_l\mathrm{D}_k - \mathrm{D}_k\mathrm{D}_l)t_{ij} = \vec{e_j}R^m_{i,kl}\vec{e_m} + \vec{e_i}R^m_{j,lk}\vec{e_m} = R^m_{i,kl}t_{mj} + R^m_{j,lk}t_{im} \tag{0.4.9}$$

Operating on the left-hand side with the differential operator D_r, we get

$$(\mathrm{D}_r\mathrm{D}_l\mathrm{D}_k - \mathrm{D}_r\mathrm{D}_k\mathrm{D}_l)t_{ij} = (\mathrm{D}_r R^m_{i,lk})t_{mj} + R^m_{i,lk}(\mathrm{D}_r t_{mj}) + (\mathrm{D}_r R^m_{j,lk})t_{im} + R^m_{j,lk}(\mathrm{D}_r t_{im})$$

In (0.4.9), we assume t_{ij} to be a tensor equal to the covariant derivative of $\vec{e_i}$. Thus, $t_{ij} = \mathrm{D}_j\vec{e_i}$, and (0.4.9) may be written as

$$(\mathrm{D}_l\mathrm{D}_k - \mathrm{D}_k\mathrm{D}_l)\mathrm{D}_j\vec{e_i} = R^m_{i,kl}(\mathrm{D}_j\vec{e_m}) + R^m_{j,lk}(\mathrm{D}_m\vec{e_i}) \tag{0.4.10}$$

We may eliminate the second term by cyclic permutation of j, l, k, by summing all three equations, and then by using the property that

$$R^m_{j,lk} + R^m_{k,jl} + R^m_{l,kj} = 0 \tag{0.4.11}$$

The left-hand side of (0.4.11) is identical to (0.4.8) if we change r by j in (0.4.8). So, the left-hand side of (0.4.11) becomes

$$\begin{aligned} &\mathrm{D}_j(R^m_{i,kl})\vec{e_m} + R^m_{i,lk}(\mathrm{D}_j\vec{e_m}) + \text{cyc. perm. of } j,l,k \\ &\qquad = R^m_{i,lk}(\mathrm{D}_j\vec{e_m} + \text{cyc. perm. of } j,l,k \end{aligned} \tag{0.4.12}$$

The second term of (0.4.12) plus its corresponding cyclic permutation cancel with the term on the right side of (0.4.12). What remains is

$$\mathrm{D}_j(R^m_{i,kl})\vec{e_m} + \text{cyc. perm. of } j,l,k$$

Since $\vec{e_m}$ appears in all the terms, then

$$\mathrm{D}_r R^j_{i,lm} + \mathrm{D}_l R^j_{i,mr} + \mathrm{D}_m R^j_{i,rl} = 0 \tag{0.4.13}$$

This is called the Bianchi Identity. Thus, the curvature tensor $R^m_{i,kl}$ fulfills the identity of (0.4.11) as well as that of (0.4.13). This is a third order derivative of the vector $\vec{e_i}$, whose variation gives rise to the curvature tensor.

By substituting $m = j$ and contracting the index i with l, we obtain

$$2\mathrm{D}_l R^l_r - \mathrm{D}_r R = 0 \ ,$$

which becomes, after using the Ricci Theorem,

$$\mathrm{D}_l(R^l_r - \frac{1}{2} g^l_r R) = 0 \ .$$

By introducing the symmetric tensor S_{rl},

$$S_{rl} = R_{rl} - \frac{1}{2} g_{rl} R \ ,$$

we obtain a conservative equation for this tensor:

$$\mathrm{D}_l(S_{rl}) = 0 \ .$$

This equation plays an important role in the Einstein interpretation of the gravitational fields of General Relativity.

Einstein looked into the metric tensor in the 4-dimensional Riemann Space, g_{ij}, which is a function of q^0, q^1, q^2, q^3, as the source of the gravitational potential. He proposed an equation similar to the Laplace and Poisson Equations which expresses the conservation of gravitational energy. Thus, the tensor of the second order, say $S_{\lambda\mu}$, that Einstein uses as representing the gravitational energy must obey the Laplace type of equation

$$S_{\lambda\mu} = 0$$

if no other actions exists on the particle but the gravitational one. This equation implies that the second order derivatives of the metric tensor are equal to zero (see (0.4.2)). Thus, this is a type of Laplace equation for the $g_{\mu\nu}$ viewed as potential functions.

With other energies besides the gravitational, like electromagnetic or microstresses of the given fluid in question, the equation must be a Poisson type, or

$$S_{\lambda\mu} = \chi Q_{\lambda\mu} \ ,$$

where the tensor $Q_{\lambda\mu}$ represents all other actions and χ is a universal constant. Einstein uses the following expression for $S_{\lambda\mu}$:

$$S_{\lambda\mu} = R_{\lambda\mu} - \frac{1}{2} g_{\lambda\mu} R$$

which we show holds the Laplace Equation. For $Q_{\lambda\mu}$, Einstein visualized the medium as a continuous one with charges particles. Thus the energy is represented by the sum of the "mechanical" impulse - energy tensor $\mathcal{P}^{\lambda\mu}$ and the electromagnetic impulse - energy tensor $\mathcal{M}_{\lambda\mu}$:

$$\mathcal{P}^{\lambda\mu} = \rho u^{\lambda} u^{\mu} + \frac{1}{c^2} T^{\lambda\mu} \tag{0.4.14}$$

and

$$\mathcal{M}_{\lambda\mu} = \frac{1}{4\pi c^2} \left(\frac{1}{4} g_{\lambda\mu} F^{\alpha\beta} F_{\alpha\beta} - F_{\lambda\alpha} F^{\alpha}_{\mu} \right) \ . \tag{0.4.15}$$

Thus

$$Q_{\lambda\mu} = \mathcal{M}_{\lambda\mu} + \mathcal{P}_{\lambda\mu} \ . \tag{0.4.16}$$

The tensor $T^{\lambda\mu}$ represents the mechanical flow of energy transmitted to the surface of an element of volumen of the continuous medium which surrounds the point P that moves with velocity u^{λ} in the 4-dimensional space. In the expression of $\mathcal{M}_{\lambda\mu}$, $F^{\alpha\beta}$ is the electromagnetic field.

Thus, the field equation is written as

$$R_{\lambda\mu} - \frac{1}{2} g_{\lambda\mu} R = \chi Q_{\lambda\mu} \ .$$

Since $\mathrm{D}_{\lambda} S_{\lambda\mu} = 0$, then $\mathrm{D}_{\mu} Q_{\lambda\mu} = 0$, expressing the conservation law for the other fields.

Particular values of the tensor $Q_{\lambda\mu}$:

1. *No electromagnetic field and no stress in the medium*: the only energy in consideration is that of the mass motion

$$Q^{\lambda\mu} = \rho u^\lambda u^\mu \tag{0.4.17}$$

where $u^\lambda = \dfrac{\mathrm{d}q^\lambda}{\mathrm{d}s}$ and ρ is the density of the matter with respect to the frame at rest.

2. *Case where the medium is a perfect one; that is, there is no shear, only pressure P*:

$$\frac{1}{c^2}T^{\lambda\mu} = -\frac{P}{c^2}g^{\lambda\mu} + \frac{P}{c^2}u^\lambda u^\mu$$

or

$$Q^{\lambda\mu} = \Big(\rho + \frac{P}{c^2}\Big)u^\lambda u^\mu - \frac{P}{c^2}g^{\lambda\mu} \tag{0.4.18}$$

3. *Only electromagnetic field, no matter:*

$$Q_{\lambda\mu} = M_{\lambda\mu} \tag{0.4.19}$$

In the first case, $\mathrm{D}_\mu Q^{\lambda\mu} = 0$ means

$$\mathrm{D}_\mu(\rho u^\mu)u^\lambda + \rho u^\mu \mathrm{D}_\mu u^\lambda = 0$$

Since $u^\lambda = \dfrac{\mathrm{d}q^\lambda}{\partial s}$ it is ($u^\lambda u^\mu = 1$) and tangent to the curse s, whose differential is ds. This curve is by definition the geodesic of the point whose mass is ρ. By being u^λ unitary,

$$u^\mu \mathrm{D}_\mu u^\lambda = 0$$

and therefore, by calling $J = \rho u$ the "mass current,"

$$u^\lambda \mathrm{D}_\mu J^\mu = 0$$

Multiplying by $u_\lambda \rightarrow$

$$u^2 \mathrm{D}_\mu(J_\mu) = \mathrm{D}_\mu(J_\mu) = 0 \tag{0.4.20}$$

which tells us that the divergency of the current is zero.

0.5 Lecture 5

0.5.1 Weyl Space

Variation of the Length of a Vector as it Moves Parallel to Itself: Weyl Geometry

In the previous sections we found that a parallel transport of a vector in Riemann Space was associated with the nullity of the absolute or tensor derivative. For the contravariant component this means that

$$\frac{\partial v^m}{\partial q^k} + v^i \Gamma^m_{ik} = 0$$

The Christoffel symbol $\Gamma^{\cdot}_{\cdot\cdot}$ may be associated with the variation of the unit vector $\vec{e_i}$ as it moves from P to $P + \delta P$, two close points in the space. Thus,

$$\mathrm{d}\vec{e_i} = \Gamma^j_{ki} \vec{e_j} \mathrm{d}q^k$$

As we remarked before, $\Gamma^{\cdot}_{\cdot\cdot}$ is not a tensor but rather a function that links two close points, like a bridge in the space. As expressed in the previous section, Γ represents the linearization of $\mathrm{d}\vec{e_i}$ in the space.

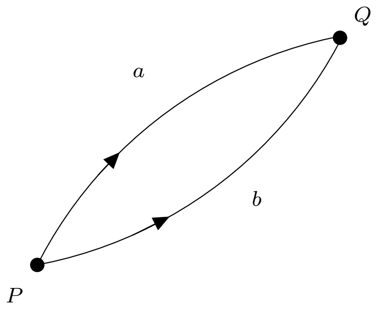
Weyl introduced another function that represents the linearization of the variation of l (the length of a vector); explicitly:

$$\mathrm{d}l = \Phi_\mu \mathrm{d}q^\mu$$

The existence of the vector function Φ_μ (a.k.a., the differential potential) implies the possibility that $l = v^\mu v_\mu$ varies from point to point. The Christoffel symbols will be different from those in the Riemann Space. They have to include the differential potential Φ_μ.

Transport of a scalar l around a closed curve

Let us assume that the scalar l moves from P to Q via the paths a and b. We will look for the variation of l as the paths vary.



The differential change of l may be written as the scalar product of two vectors:

$$\frac{l(q+\delta q)-l(q)}{l}=\frac{\mathrm{d}l}{l}=\Phi_\mu \mathrm{d}q^\mu$$

The circulation around the curve will induce a work of Φ_μ. We ssume that this vector Φ_μ moves parallel to itself.

The circulation from P to Q via a and back to P via b is given by

$$\oint_{P_aQ}\Phi_\mu \mathrm{d}q^\mu+\oint_{Q_bP}\Phi_\mu \mathrm{d}q^\mu$$

By the Stokes Theorem, the circulation may be expressed as the flux of the rotational of Φ_μ on the surface inclosed by the curve

$$\oint \Phi_\mu \mathrm{d}q^\mu=\iint(\nabla\times\Phi_\mu)\ \mathrm{d}s$$

If the rotational is zero, the function $l(q)$ is simple and there is integrability.

If the rotational of Φ_μ is different from zero, then

$$(l_a(Q)-l_a(P))+(l_b(P)-l_b(Q))\neq 0$$

In this case, l at Q is different from l at P, and thus the difference of the circulation of Φ_μ (the *differential potential*) will depend on the path. We say that Φ_μ is not integrable. That is,

$$\frac{\partial\Phi_\nu}{\partial x^\mu}-\frac{\partial\Phi_\mu}{\partial x^\nu}\neq 0$$

The scalar l may be written as the scalar product of two vectors, $\vec{u}$ and $\vec{v}$, in the space whose metric is $g_{\mu\nu}$:

$$l = u^\mu v^\nu g_{\mu\nu}$$

and so

$$l + \mathrm{d}l = (u^\mu + \mathrm{d}u^\mu)(v^\nu + \mathrm{d}v^\nu)(g_{\mu\nu} + \mathrm{d}g_{\mu\nu})$$

As we have shown previously, the Christoffel symbols are related to $g_{\mu\nu}$ and its derivative, so Φ_μ also has to be related to them. Let us see how: neglecting second-order terms and indicating with "d" the derivative of a scalar (not a tensor derivative),

$$\mathrm{d}l = (\mathrm{d}u^\mu)v^\nu g_{\mu\nu} + u^\mu \mathrm{d}v^\nu g_{\mu\nu} + u^\mu v^\nu \mathrm{d}g_{\mu\nu} \ . \tag{0.5.1}$$

Now, using $\dfrac{\mathrm{d}l}{l} = \Phi_\rho \mathrm{d}q^\rho$,

$$\mathrm{d}l = l\Phi_\rho \mathrm{d}q^\rho = u^\mu v^\nu g_{\mu\nu}\Phi_\rho \mathrm{d}q^\rho \tag{0.5.2}$$

Before equating these equations for dl, let us introduce the generalized Christoffel symbols $\overline{\Gamma}^{\cdot}_{\cdot\cdot}$ by the following:

$$\mathrm{d}u_\nu = \overline{\Gamma}^\mu_{\rho\nu}\mathrm{d}q^\rho u_\mu$$

and

$$\mathrm{d}v^\mu = -\overline{\Gamma}^\mu_{\rho\nu}\mathrm{d}q^\rho v^\nu \ .$$

These equations imply that the absolute tensor derivative $\mathrm{D}_\nu u = u_\mu v^\mu = 0$; in other words, the parallel transport of $\vec{u}$ and $\vec{v}$.

Since dl is a scalar, $\mathrm{d}l = \dfrac{\partial l}{\partial q}\mathrm{d}q$ and

$$\mathrm{d}l = v^\nu g_{\mu\nu}\mathrm{d}u^\mu + u^\mu g_{\mu\nu}\mathrm{d}v^\nu + v^\mu u^\mu \mathrm{d}g_{\mu\nu} \ .$$

From the previous definition of $\overline{\Gamma}$,

$$\mathrm{d}_\rho v^\nu = -\overline{\Gamma}^\nu_{\alpha\rho} v^\alpha \mathrm{d}q^\rho \quad \text{and} \quad \mathrm{d}_\rho u^\mu = -\overline{\Gamma}^\mu_{\rho\beta} u^\beta \mathrm{d}q^\rho$$

so

$$\mathrm{d}l = -v^\nu g_{\mu\nu}\overline{\Gamma}^\mu_{\rho\beta}u^\beta \mathrm{d}q^\rho - g_{\mu\nu}u^\mu\overline{\Gamma}^\nu_{\alpha\rho}v^\alpha \mathrm{d}q^\rho + v^\nu u^\mu \mathrm{d}_\rho(g_{\mu\nu})$$

On the other hand, by linearizing $\mathrm{d}l$ as $\dfrac{\mathrm{d}l}{l} = \Phi(q)\mathrm{d}q$,

$$\mathrm{d}l = l\mathrm{d}q^k\Phi_k = v^\nu u^\mu g_{\mu\nu}\Phi_k \mathrm{d}q^k$$

Comparing these two equalities for $\mathrm{d}l$,

$$-v^\nu u^\beta g_{\mu\nu}\overline{\Gamma}^\mu_{\rho\beta}\mathrm{d}q^\rho - g_{\mu\nu}u^\mu v^\alpha\overline{\Gamma}^\nu_{\alpha\rho}\mathrm{d}q^\rho + v^\nu u^\mu \mathrm{d}_\rho(g_{\mu\nu}) = v^\nu u^\mu g_{\mu\nu}\Phi_k \mathrm{d}q^k$$

In the first two terms, we use the symbols $\Gamma_{\dots}$ as

$$g_{\mu\nu}\overline{\Gamma}^\mu_{\rho\beta} = \overline{\Gamma}_{\rho\nu\beta} \quad \text{and} \quad g_{\mu\nu}\overline{\Gamma}^\nu_{\alpha\rho} = \overline{\Gamma}_{\alpha\mu\rho}$$

Substituting into the previous equation, changing k by ρ in the second member, and then dividing by $\mathrm{d}q^\rho$:

$$-v^\nu u^\beta\overline{\Gamma}_{\rho\nu\beta} - u^\mu v^\alpha\overline{\Gamma}_{\alpha\mu\rho} + v^\nu u^\mu\frac{\partial g_{\mu\nu}}{\partial q^\rho} = v^\nu u^\mu g_{\mu\nu}\Phi_\rho$$

Now by changing β by μ and α by ν in the first two terms (they are repeated indices) and dropping $v^\nu u^\mu$, we get:

$$-\overline{\Gamma}_{\rho\nu\mu} - \overline{\Gamma}_{\nu\mu\rho} + \frac{\partial g_{\mu\nu}}{\partial q^\rho} = g_{\mu\nu}\Phi_\rho \tag{0.5.3}$$

If $\Phi_\rho = 0$, this equation is the same as that which relates the original Christoffel symbols to the variation of the metric tensor. In the case $\mathrm{d}l = 0$, that is, the scalar vector of $\vec{u}$ and $\vec{v}$ as they move parallel to themselves does not change. If $\Phi_\rho \neq 0$, the generalized Christoffel symbols $\overline{\Gamma}$ will depend on the variation of the metric tensor and the differential potential Φ_ρ. Soon, we will see that the metric tensor may change with the gauge, but $\overline{\Gamma}$ will remain constant.

Generalization of the Ricci Theorem

Let us differentiate the length l of a vector $\vec{\xi}$, such that $l = \xi^i \xi^j g_{ij}$:

$$\mathrm{d}_m l = (\mathrm{D}_m \xi^i)\xi^j g_{ij} + \xi^i g_{ij}(\mathrm{D}_m \xi^j) + \xi^i \xi^j \mathrm{D}_m(g_{ij})$$

By definition of $\overline{\Gamma}$, and assuming that the vector $\vec{\xi}$ moves parallel to itself,

$$\mathrm{D}_m \xi^i = \frac{\partial \xi^i}{\partial q^m} + \overline{\Gamma}^i_{ml}\xi^l = 0$$

and similarly for $\mathrm{D}_m \xi^j$. So, $\mathrm{d}_m l = \xi^i \xi^j \mathrm{D}_m(g_{ij})$ is zero if $\mathrm{D}_m(g_{ij})$ is zero.

Now, taking g_{ij} as a tensor of the second order,

$$\mathrm{D}_m(g_{ij}) = \frac{\partial g_{ij}}{\partial q^m} - \overline{\Gamma}^l_{mi} g_{jl} - \overline{\Gamma}^k_{mj} g_{ik}$$

and according to (0.5.3),

$$\mathrm{D}_m(g_{ij}) = g_{ij}\Phi_m$$

We may say that in the Weyl picture, this is a generalization of the Ricci Theorem (for parallel transport only, but Ricci is general in the Riemann Space).

Invariance of the Angle Between Two Vectors as They Move Parallel to Themselves

From the previous equation, we have found a useful assertion: the absolute or total derivative D of the covariant component of $\vec{\xi}$ is not zero as $\mathrm{D}\xi^i = 0$, but rather

$$\mathrm{D}_m \xi_i = \mathrm{D}_m(\xi^j g_{ij}) = (\mathrm{D}_m \xi^j)(g_{ij}) + \xi^j(\mathrm{D}_m g_{ij})$$

The first term is zero, and for the second we use the generalized Ricci Equation:

$$\mathrm{D}_m \xi_i = \xi^j(\mathrm{D}_m g_{ij}) = \xi^j \Phi_m g_{ij}$$

and thus the absolute derivative of the covariant component is

$$\mathrm{D}_m \xi_i - \xi^j \Phi_m g_{ij} = 0$$

Let us apply this theorem to find the angle between two vectors in parallel displacement:

$$\cos\theta = \frac{\xi_i \eta^i}{\sqrt{l_1 l_2}}$$

where l_1 and l_2 are the lengths of ξ and η. The total derivative of $\cos\theta$, which is a scalar, is

$$\mathrm{D}_m \cos\theta = \frac{(\mathrm{D}_m \xi_i)\eta^i}{\sqrt{l_1 l_2}} + \frac{\xi_i (\mathrm{D}_m \eta^i)}{\sqrt{l_1 l_2}} + \xi_i \eta^i \mathrm{D}_m \frac{1}{\sqrt{l_1 l_2}}$$

The second terms is zero. The first is

$$\frac{(\mathrm{D}_m \xi_i)\eta^i}{\sqrt{l_1 l_2}} = \frac{\eta^i \xi^j \Phi_m g_{ij}}{\sqrt{l_1 l_2}}$$

The third term is

$$\xi_i \eta^i \mathrm{D}_m \frac{1}{\sqrt{l_1 l_2}} = \eta^i \xi_i \Big(-\frac{1}{2}\Big) \frac{(\mathrm{d}l_1) l_2 + (\mathrm{d}l_2) l_1}{\sqrt{(l_1 l_2)^3}}$$

Since $\mathrm{d}_m l_1 = l_1 \dfrac{\mathrm{d}\Phi}{\mathrm{d}q^m} = l_1 \Phi_m$ and $\mathrm{d}_m l_2 = l_2 \Phi_m$, where $\mathrm{d}_m l = \dfrac{\mathrm{d}l}{\mathrm{d}q^m}$, etc., the third term becomes

$$\eta^i \xi_i \Big(-\frac{1}{2}\Big) \frac{(\mathrm{d}l_1) l_2 + (\mathrm{d}l_2) l_1}{(l_1 l_2)\sqrt{l_1 l_2}} = -\frac{\xi_i \eta^i \Phi_m}{\sqrt{l_1 l_2}} = -\frac{\eta^i \xi^j g_{ij} \Phi_m}{\sqrt{l_1 l_2}}$$

which cancels with the first term. Thus, $\mathrm{D}_m \cos\theta = 0$, and θ is invariant.

Relationship Between Weyl $\bar{\Gamma}$ and the Christoffel Symbols

$$\frac{\partial}{\partial q^k} g_{ij} = g_{ij} \Phi_k + \overline{\Gamma}_{kji} + \overline{\Gamma}_{jik}$$

and similarly,

$$\frac{\partial}{\partial q^j} g_{ki} = g_{ki}\Phi_j + \overline{\Gamma}_{jki} + \overline{\Gamma}_{jik} \quad \text{and} \quad \frac{\partial}{\partial q^i} g_{kj} = g_{jk}\Phi_i + \overline{\Gamma}_{jki} + \overline{\Gamma}_{kji}$$

Summing these three equations and dividing by 2,

$$\frac{(\partial_k g_{ij} + \partial_j g_{ki} - \partial_i g_{kj})}{2} = \frac{1}{2}(g_{ij}\Phi_k + g_{ki}\Phi_j - g_{jk}\Phi_i) + \frac{\overline{\Gamma}_{jik} + \overline{\Gamma}_{jik}}{2}$$

The first term is, in the Riemann Space, the value of Γ_{jik} of the old Christoffel symbols. Thus, the new Christoffel symbols and the old ones are related by

$$\overline{\Gamma}_{jik} = \Gamma_{jik} - \frac{1}{2}(g_{ij}\Phi_k + g_{ki}\Phi_j - g_{jk}\Phi_i)$$

Curvature and Generalized Christoffel Symbols

Let us define a vector $\vec{v}$ that moves parallel to itself in a closed curve, that may lie on a curved surface, and such that it is related to the connecting function by

$$\mathrm{d}v^i = \Phi^i_k \mathrm{d}q^k$$

We add i to the set of Φ_k to identify in it the given contravariant component of $\vec{v}$. Since it moves parallel to itself,

$$\mathrm{d}v^i = \overline{\Gamma}^i_{jk} v^j \mathrm{d}q^k$$

The circulation of $\vec{v}$ on the closed curve will, in general, be different from zero. So the final variation of $\vec{v}$ is

$$\mathrm{d}v^i = \oint \Phi^i_k \mathrm{d}q^k = \frac{1}{2}\iint_s (\Phi^i_{k,l} - \Phi^i_{l,k})\mathrm{d}s^{lk}$$

This variation will be, as done before in the Riemann Space, proportional to the curvature of the surface:

$$\mathrm{d}v^i = -\frac{1}{2}\iint R^i_{j,kl} v^j \mathrm{d}s^{lk}$$

where

$$R^i_{j,kl} = \frac{\partial}{\partial q^k}\overline{\Gamma}^i_{jl} - \frac{\partial}{\partial q^l}\overline{\Gamma}^i_{jk} + \overline{\Gamma}^i_{mk}\overline{\Gamma}^m_{jl} - \overline{\Gamma}^i_{ml}\overline{\Gamma}^m_{jk}$$

0.6 Lecture 6

0.6.1 Gauge Theory

Introducing Different Gauges, or "Etalon," in the Same Space

As we will see in this section Weyl assumed the possibility of changing the length of the vector by changing the unit of measure, or "etalon" in French. This will correspond to two different Riemann spaces superimposed in one. Thus, when a vector moves from one point P to a closed one Q while remaining parallel to itself, the metric tensor $g_{ij}(q)$ changes because the coordinates q change. This change corresponds, as we have seen in the previous section, to the curvature of the space. But if we change the "etalon" this brings a new reason for the variation of $g_{ij}(q)$. If g_{ij} changes by a factor λ as the "etalon" changes, then $g_{ij} \to g'_{ij} = \lambda g_{ij}$. If $l(q)$ is the length of a vector $\vec{\xi}$ as it moves parallel to itself $l(q)$ changes to $l'(q) = \lambda l(q)$. So the connecting function that we have called $\Phi = \dfrac{\mathrm{d}l}{l}$ in the previous section and which is directly related to the generalized Christoffel symbols $\overline{\Gamma}$ also changes because

$$\mathrm{d}l' = \lambda \mathrm{d}l + l\mathrm{d}\lambda$$

and

$$\Phi' = \frac{\mathrm{d}l'}{l'} = \frac{\mathrm{d}l}{l} + \frac{\mathrm{d}\lambda}{\lambda} = \Phi + \frac{\mathrm{d}\lambda}{\lambda}$$

Thus the question regarding this general parallel displacement discussed by Weyl is how the whole approach changes as $g_{\mu\nu}$ is multiplied by a positive factor λ, which may depend on the coordinates.

The element of space at the point $\mathrm{d}s^2$ changes to $\mathrm{d}s^2 = \lambda g_{\mu\nu}\mathrm{d}q^\nu \mathrm{d}q^\mu$ [redacted] The factor λ is an indication of a change in the length of the unit of measurement, "etalon". It is as if the ruler was changing its "length" as we go from one point to the next, like it was made of metal and at each point of space-time the temperature was different, causing dilation and contraction to occur. (The fact that the physics does not change as we change $g \to \lambda g$ is obvious for the light transport from a point P to $P + \mathrm{d}P$ where $\mathrm{d}s^2$ is always equal to zero.)

Let us discuss gauge change from another point of view. Let us call E the "etalon" at P on the geodesic path: with its help we measure the length of $\vec{e_i}$. Now call E' the "etalon" at P on the nongeodesic path: with its help we measure the length of $\vec{e_i}'$. E is called the "geodesic etalon" and is different for different points in the space. Thus, $|\vec{e_i}'| = |\vec{e_i}|\frac{E'}{E}$ and $g'_{ij} = \vec{e_i}' \cdot \vec{e_j}' = (\vec{e_i} \cdot \vec{e_j})\left(\frac{E'}{E}\right)^2$. If we say that $\lambda' = \left(\frac{E'}{E}\right)^2$, then $g_{ij} \to g_{ij}\lambda'$ and the length of $\vec{\xi}$ changes from $l = \xi^i \xi^j g_{ij} \to \lambda' \xi'^i \xi'^j g_{ij} = \lambda' l'$. When we move from P to Q, $\vec{\xi}$ will conserve its components, and of course its length, as long as the "etalon" is invariant. That is,

$$l(Q) = l(P) \quad \text{and} \quad \xi^i(Q) = \xi^i(P)$$

When λ is assumed to be a function of the point in the space-time, from P tp Q, then λ will vary by $\mathrm{d}\lambda$. On the other hand, by changing the "etalon" from E to E' at P itself, the length will vary at P by a factor of $\left(\frac{E'}{E}\right)^2 = \lambda$; that is

$$l' = l\lambda \quad \text{at} \quad P$$

Then, by moving to Q (a close point to P), l' will vary by

$$\mathrm{d}l' = l\mathrm{d}\lambda + \lambda \mathrm{d}l$$

so

$$\frac{\mathrm{d}l'}{l'} = \frac{l}{l\lambda}\mathrm{d}\lambda + \lambda\frac{\mathrm{d}l}{l\lambda}$$

and $\frac{\mathrm{d}l'}{l'} = \frac{\mathrm{d}l}{l} + \frac{\mathrm{d}\lambda}{\lambda}$ is the variation of l' when $P \to Q$ per unit of l'.

As in the previous section, let us call $\frac{\mathrm{d}l}{l} = \mathrm{d}\Phi$, where Φ is assumed to be a continuous function. As P moves to Q, this function varies as

$$\mathrm{d}\Phi' = \mathrm{d}\Phi + \frac{\mathrm{d}\lambda}{\lambda} = \mathrm{d}\Phi + \mathrm{d}(\log\lambda)$$

The vector Φ_μ may be seen as a field that is generated as the ”etalon” varies in the space-time worlds. Thus, in a parallel displacement, if g_{ij} changes to λg_{ij}, the length changes and at the same time the differential of the connecting function $\mathrm{d}\Phi$ is changed by the differential $\mathrm{d}(\log\lambda)$. The differential function $\mathrm{d}\Phi$ may be expressed in the generalized coordinates by its components Φ_μ, such that

$$\mathrm{d}\Phi = \Phi_\mu \mathrm{d}q^\mu$$

If the circulation of Φ_μ around a closed path is not null, then the rotational of Φ_μ generates a new field which is independent of its coordinate system, that is, invariant under coordinate transformation. Φ_μ depends on the change of the ”etalon” as well as the transport of $\vec{\xi}$ from P to Q. Thus, the change of ”etalon” changes the quadratic form $\mathrm{d}s^2$, but at the same time, in a parallel transport the differential connection function $\mathrm{d}\Phi'$ changes by the term $\dfrac{\mathrm{d}\lambda}{\lambda}$.

The equation $\mathrm{d}\Phi' = \mathrm{d}\Phi + \dfrac{\mathrm{d}\lambda}{\lambda} = \mathrm{d}\Phi + \mathrm{d}(\log\lambda)$ may be rewritten as

$$\Phi'_\mu \mathrm{d}q^\mu = \Phi_\mu \mathrm{d}q^\mu + \mathrm{d}(\log\lambda)$$

or

$$\Phi'_\mu = \Phi_\mu \frac{\mathrm{d}(\log\lambda)}{\mathrm{d}q^\mu} \tag{0.6.1}$$

If instead of λ we use e^λ, then this equation becomes

$$\Phi'_\mu = \Phi_\mu + \frac{\partial\lambda}{\partial q^\mu}$$

Introducing the variation of g, we see that

$$\mathrm{d}\Phi' = \frac{\mathrm{d}g_{ij}}{g_{ij}} + \frac{\mathrm{d}\lambda}{\lambda}$$

and

$$\Phi'_\mu = \frac{1}{g}\frac{\partial g'_{ij}}{\partial q^\mu} + \frac{\partial(\log\lambda)}{\partial q^\mu}$$

Next we will show that $\overline{\Gamma}^{\mu}_{\nu\rho}$ are invariant under changes of "etalon." [redacted] Before, we found that $g_{\mu\nu}$, $\Gamma_{\mu\nu\rho}$, and the potential Φ_ρ are linked during parallel transport of a vector by the expression

$$\frac{\partial g_{\mu\nu}}{\partial q^\rho} = g_{\mu\nu}\Phi_\rho + \overline{\Gamma}_{\rho\nu\mu} + \overline{\Gamma}_{\nu\mu\rho} \tag{0.6.2}$$

Now, if the "etalon" changes, then

$$g_{\mu\nu} \to g'_{\mu\nu} = \lambda g_{\mu\nu}$$

and

$$\Phi_\rho \to \Phi'_\rho = \Phi_\rho + \frac{1}{\lambda}\frac{\mathrm{d}\lambda}{q^\rho}$$

Then (0.6.2) changes to

$$\lambda\frac{\partial g_{\mu\nu}}{\partial q^\rho} + g_{\mu\nu}\frac{\mathrm{d}\lambda}{q^\rho} = \lambda g_{\mu\nu}(\Phi_\rho + \frac{1}{\lambda}\frac{\mathrm{d}\lambda}{q^\rho}) + \overline{\Gamma}'_{\rho\nu\mu} + \overline{\Gamma}'_{\nu\mu\rho}$$

or

$$\lambda\frac{\partial g_{\mu\nu}}{\partial q^\rho} = \lambda g_{\mu\nu}\Phi_\rho + \overline{\Gamma}'_{\rho\nu\mu} + \overline{\Gamma}'_{\nu\mu\rho}$$

Dividing by λ,

$$\frac{\partial g_{\mu\nu}}{\partial q^\rho} = g_{\mu\nu}\Phi_\rho + \frac{\overline{\Gamma}'_{\rho\nu\mu} + \overline{\Gamma}'_{\nu\mu\rho}}{\lambda} \tag{0.6.3}$$

Substituting $\overline{\Gamma}'_{\mu\nu\rho} \to \overline{\Gamma}'^{k}_{\mu\rho} g'_{k\nu} = \overline{\Gamma}'^{k}_{\mu\rho} g_{k\nu}$, and since $\overline{\Gamma}_{klj} = \overline{\Gamma}^{i}_{kj} g_{il}$, the last term of (0.6.3) becomes

$$\frac{1}{\lambda}(\overline{\Gamma}'^{k}_{\rho\mu} g_{k\nu} + \overline{\Gamma}'^{k}_{\nu\rho} g_{k\mu})$$

Similarly, the last term of (0.6.2) becomes

$$(\overline{\Gamma}^{k}_{\rho\mu} g_{k\nu} + \overline{\Gamma}^{k}_{\nu\rho} g_{k\mu})$$

Comparing (0.6.2) and (0.6.3), they are identical if $\overline{\Gamma}'^{k}_{\nu\rho} = \overline{\Gamma}^{k}_{\nu\rho}$, etc. So, $\overline{\Gamma}^{\cdot}_{\cdot\cdot}$ is invariant under gauge change.

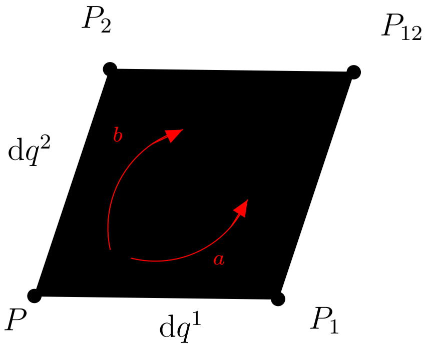
Invariance of the Absolute Derivative of ξ^i

$$\mathrm{D}_m \xi^l = \frac{\partial \xi^l}{\partial q^m} + \overline{\Gamma}^l_{mi} \xi^i \ .$$

We have seen that $g_{ij} \to g_{ij}\lambda$ leaves $\overline{\Gamma}^l_{mi}$ invariant. Thus, if the coordinates do not change (only the gauge), then $\mathrm{D}_m \xi^l$ remains constant and in parallel transport is equal to zero.

Invariance of Curvature as Gauge Changes

Let us show that the curvature tensor defined from the geometric idea of parallel transport is invariant under gauge change.



In the above infinitesimal parallelogram, we may start at point P with our transport of a vector $\vec{v}$ and reach P_{12} via the path a or b. At P_{12}, we compare the two vectors: in general, they will not coincide except if the parallelogram is on a plane. We take the difference between the two vectors as the element proportional to the curvature:

$$\vec{v}_a(P_{12}) - \vec{v}_b(P_{12}) = \Delta\vec{v}$$

Then, if $\Delta\vec{v}$ may be written as the product of the curvature tensor per unit of surface ΔR^μ_ν, such that

$$\Delta v^\mu = \Delta R^\mu_\nu v^\nu$$

In general, ΔR^μ_ν is proportional to the surface of the parallelogram, and so

$$\Delta R^\mu_\nu = R^\mu_{\nu\rho\sigma}\mathrm{d}q^\rho \mathrm{d}q^\sigma$$

Since the contravariant components of the vector $\vec{v}$, v^μ and Δv^μ, do not change under a gauge change, then R^μ_ν and $R^\mu_{\nu\rho\sigma}$ are also invariant.

The tensor $R^\mu_{\nu\rho\sigma}$ is antisymmetric in the interchange of the indices ρ and σ. This is so because the element $(\Delta q)^{\sigma\rho} = \mathrm{d}q^\sigma \mathrm{d}q^\rho$ refers to an oriented surface, and so changes sign when it goes counterclockwise rather than clockwise. Thus,

$$\Delta q^{\sigma\rho} = -\Delta q^{\rho\sigma}$$

and

$$R^{\mu}_{\nu\rho\sigma} = -R^{\mu}_{\nu\sigma\rho}$$

Thus, if we use the subscripts ρ and σ only once and ignore the order, then

$$\Delta R^{\mu}_{\nu} = \frac{1}{2} R^{\mu}_{\nu\rho\sigma} \Delta q^{\rho\sigma}$$

The curvature tensor $\overline{R}^{\mu}_{\nu\rho\sigma}$ is, in this approach, generalized from the gravitational field $P^{\mu}_{\nu\rho\sigma}$ as well as from the electromagnetic field vector $F_{\rho\sigma}$. It may be seen that the additive contribution of both fields may be expressed as

$$\overline{R}^{\mu}_{\nu\rho\sigma} = P^{\mu}_{\nu\rho\sigma} - \frac{1}{2}\delta^{\mu}_{\nu} F_{\rho\sigma}$$

where δ^{μ}_{ν} is the Kronecker Function: 1 or 0. The tensor $P^{\mu}_{\nu\rho\sigma}$ is antisymmetric in both pairs: μ, ν and ρ, σ.

Obviously, when the electromagnetic field is not present, $\overline{R}^{\mu}_{\nu\rho\sigma}$ corresponds to the classical Ricci Tensor and parallel displacement preserves the length of the vector. However, when $F_{\rho\sigma}$ is introduced, we do not have a unique geodesic and the integral will depend on the path taken. In any case, the change in gauge or ”etalon” does not change the curvature tensor.

Appendix

1. Some Remarks as to the Differences Between Euclidean, [redacted] Riemann, and Weyl Geometries

A space is an object that becomes expressed when a set of numbers is associated to it, and consequently becomes a nest of functions. Thus, if the set R of real numbers is associated with a 3D space, then we have the usual E^3 space, and a point in the space is a set of real number (x, y, z). A function of $f(x, y, z)$ will be defined as a mapping of a region A of the space, which corresponds to a subset of R that goes over to another region B of R, which is occupised by the numbers $r_1 = f(x_1, y_1, z_1), r_2 = f(x_2, y_2, z_2)$, etc. This mapping is expressed as

$$f : A \rightarrow E^3 \quad , \quad x, y, z, \epsilon A$$

A very close definition to this is that of a *patch*, a concept which is important in topology. In this case, a 2D subspace, described by the two real numbers (x, y), is mapped over a space E^3. This final mapping is called a surface in E^3. It is expressed similarly as

$$f : A \rightarrow E^3$$

These definitions are independent of the space itself, and may be used in any other space.

A space with geometry in it is a space where distances and angles may be measured. Because of this, it is traditional to use the dot product as the fundamental tool. Euclidean Geometry is built with the dot product between two vectors:

$$\vec{v_1} = (x_1, y_1, z_1) \quad \text{and} \quad \vec{v_2} = (x_2, y_2, z_2)$$

$$\vec{v_1} \cdot \vec{v_2} = g_{ij} x^i x^j$$

The Euclidean Geometry preserves the elements of the space as curves, triangles, surfaces, etc. Those that are congruent to themselves are *isometries*, or mappings that preserve distances between points. That is, if F is the mapping and d indicates distance,

$$f : E^3 \to E^3$$

$$d(F(p), F(q)) = d(p, q)$$

for all points p and q of E^3. Translation, rotation, etc. or a rigid body are examples of these isometries. The figures are called congruents because one can be superimposed on the other, and they are identical although located at different parts of the space. Among the properties preserved are the geodesic and the curvature; both concepts were used by Einstein in the Minkowsky Space of four dimensions in order to express the gravitational laws in General Relativity.

2. **Riemann Space and Riemann Geometry**

As in the Euclidean Geometry, the concept of elements is preserved by isometries. From a point of view of differentiability, the Riemann Space can be visualized as an infinite number of Euclidean Spaces; that is, a space where each point point has a neighborhood that may be taken as a Euclidean Space. In addition, the passage from on region to the next is smooth, so differentiability exists in the local region as well as in the overlapping of these regions. Another way to express this fact is by referring to Riemann space as a Compact C^∞ Manifold with a metric g_{ij} that is greater than zero, differentiable, symmetric, and bilinear. "C^∞" expresses the fact that it is infinitely differentiable.

In the four-dimensional Riemann Space, the geodesics are also world curves which represent the history of the particle in consideration. Those curves are tangent to the velocity vector in 4D, and they are also the stationary path between two points.

The algebra of the Riemann Space V_n is carries out in the tangent or *osculatrix* plane. At each point of V_n, we create this plane where the tensor algebra brings the same results as if we had used the same narrow interval around the point of contact of the space. If we use a set of coordinates $y^1, y^2, y^3, \ldots, y^n$ at the point M of V_n, then there is a point m (the contact point) in the tangent plane that follows a Euclidean algebra such that

(a) The fundamental tensor $g_{\mu\nu}$ at M in V_n is the same g_{ij} at the point m in the Euclidean plane ϵ_n.

(b) The first order variations of these metric tensors, when we move from M to $M+\Delta M$ and therefore from m to $m+\Delta m$, are the same up to the second order in the variation of y^i. That is, we neglect $(y^i(m) - y^i(m + \Delta m))$.

(c) The coordinates y^μ in V_n at M become y^i at m in the Euclidean tangent space. (We use the Greek scripts for the Riemann Space and the Latin scripts for the Euclidean Space.) The y^i in ϵ_n become curvilinear coordinates.

Thus, according to (a) and (b), we assume

$$(g_{ij})_m = (g_{\mu\nu})_M$$

and

$$\left(\frac{\mathrm{d}g_{ij}}{\mathrm{d}y^k}\right)_m = \left(\frac{\mathrm{d}g_{\mu\nu}}{\mathrm{d}y^k}\right)_M$$

This assumes that the Christoffel symbols are the same at M and m, and also that the parallel transport of a vector M to ΔM agrees with the parallel transport of a vector from m to Δm in ϵ_n. Thus,

$$(\mathrm{D}_k v^\mu)_M = (\mathrm{D}_k v^i)_m = \frac{\partial v^i}{\partial y_k} + \Gamma^i_{jk} v^j$$

If the point M describes a curve C in V_n, there is an image curve γ in ϵ_n, where at every point M $g_{\mu\nu}(C)$ is equal to $g_{ij}(\gamma)$ at m over γ. This implies a multiple of Euclidean planes at the different point M of C.

The fundamental portion of arc at C is the same at M and m. That is,

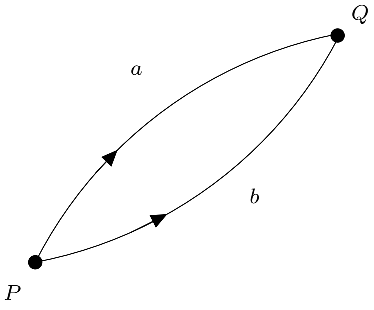
$$ds^2 = g_{\mu\nu}(M)dy^\mu dy^\nu = g_{ij}(m)dy^i dy^j$$

The curvature of the Riemann Space at M in C is found from the Christoffel symbols at m going from m to Δm in the Euclidean Space ϵ_n. This assumes that the second derivative of the fundamental tensor is the same at M in V_n and m in ϵ_n.

3. **Summary of Weyl Geometry Approach to Gauge Change** [redacted]

We start by scaling the metric tensor g_{ij} by a factor of λ, such that $g'_{ij} = \lambda g_{ij}$. So, $l' = \lambda l$, where l is the square of a vector $\vec{v} \to dl =$ work of field $= \Phi_\mu dq^\mu$. In order to preserve, during a parallel displacement which yields a change in the magnitude of the vector, the same field values and the same geodesic and curvature, the variation of l is written as the work of a potential field Φ_μ that does not affect the original field if it always expresses the change per unit length of $\frac{dl}{l}$.

Weyl Space is a generalization of Riemann Space. Now, between two points in the space there is not a unique geodesic. Taking a vector $\vec{v}$ at P a moving it parallel to itself, we reach the point Q via path a.



This path is a geodesic where, if τ is a parameter describing the vector $v^\mu = \frac{dq^\mu}{d\tau}$ tangent to it, the equation for the geodesic is such that the absolute derivative $\mathrm{D}_\tau v^\mu = 0$, or

$$\frac{\mathrm{d}}{\mathrm{d}\tau}\frac{\mathrm{d}q^\mu}{\mathrm{d}\tau} + \Gamma^\mu_{\rho\nu} = 0$$

or

$$\frac{\mathrm{d}^2 q^\mu}{\mathrm{d}\tau^2} + \Gamma^\mu_{\rho\nu}\frac{\mathrm{d}q^\nu}{\mathrm{d}\tau}\frac{\mathrm{d}q^\rho}{\mathrm{d}\tau} = 0$$

But if we now allow the length of the vector V^μ to vary, we may choose another path b that goes from P to Q and still have a geodesic which holds similar equations. In this case, the integration of $v^\mu \cdot V_\mu$ around the loop $a + b$ starting at point P and returning to P is different from zero and equal to the work of a function Φ which in principal is different for each loop and is part of the new Christoffel symbols. Next, as we change $g_{\mu\nu}$ to $\lambda g_{\mu\nu}$, some tensors may vary, while others may not. We have seen that the curvature tensors as well as the electromagnetic tensor do not change under this gauge change.

In the development of Weyl, the parallel displacement of a vector as that given by the absolute derivative of the contravariant components is again equal to zero, such that $\mathrm{D}v^\mu = 0$. But now, the Γ symbols are not equal to the Christoffel symbols Γ, but instead a similar equation holds, namely,

$$\frac{\partial v_\mu}{\partial q^\rho} - \Gamma^\alpha_{\rho\mu} v_\alpha = 0$$

for the covariant, and

$$\frac{\mathrm{d}v^\mu}{\mathrm{d}q^\rho} + \Gamma^\mu_{\rho\alpha} v^\alpha = 0$$

for the contravariant components of vector $\vec{v}$. This is a simple extension of Riemann Geometry, but $\Gamma^\mu_{\rho\alpha}$ are not only functions of the metric tensor g_{ij}, as before, but also of the potential function Φ.

In order to preserve the ”connecting” function $\Gamma^\mu_{\rho\alpha}$ in the above formula as we change the gauge or etalon, Φ_μ must vary with the change of etalon as follows: if $g_{\mu\nu} \to \lambda g_{\mu\nu}$, then

$$\Phi_\mu \to \Phi_\mu + \frac{\partial \log \lambda}{\partial q^\mu}$$

This will preserve the objects in the 4D geometry, like the curvature and the circulation of a vector parallel to itself, which eventually will be used to express the physical laws.

According to Mie, all the laws of nature correspond to an integral of the "action density," which must be specified for each phenomenon. This integral must be stationary as the potentials (or state functions) of the system vary. It is assumed that the 4D domain where the law applied is surrounded by a surface where the potentials vanish. The action is a simple scalar expressed as

$$\text{Action Density} = S = \int W \ \mathrm{d}\Omega$$

where Ω is the element of volumen in the space and W is the action density. Since $\mathrm{d}\Omega \propto \mathrm{d}s^2 \approx g_{ij} g_{kl} \mathrm{d}x^i \mathrm{d}x^k \mathrm{d}x^l$ and $g_{\mu\nu} \to \lambda g_{\mu\nu}$, then $\mathrm{d}\Omega$ must change proportionally to λ^2 and Q must change proportionally to λ^{-2} in order for S to be invariant.

We are going to pursue this line of work in future chapters. But before we leave this discussion, we would like to outline a parallel between the Weyl approach (or the gauge variation established earlier) and the Theory of Electromagnetism. In classical Maxwell Theory, the relevant objects are the field strength E, B, and the charge e. However the four potentials vectors $(\Phi_0, \Phi_1, \Phi_2, \Phi_3)$ are introduced to make the theory mathematically sound.

0.7 Lecture 7

0.7.1 Analytic Dynamics

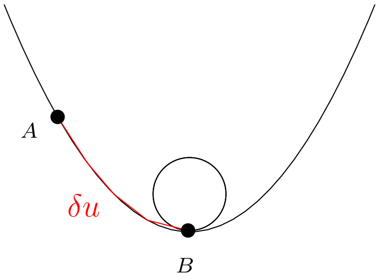
Variational Calculus and the Euler Differential Equation

The extreme of a function belongs to Ordinary Calculus. The extreme of an integral belongs to Variational Calculus.

Virtual changes of a variable of a function, which represents a real situation, are different from the actual physical variation.

For example, a ball at the bottom of a well with no external forces but gravity acting on it does not move, but it may suffer a virtual displacement δu from B to A and we see it as a mathematical variation which implies a variation on the function $F(u)$ such that

$$\delta F = \frac{\partial F}{\partial u}\delta u \ .$$



The following is an example of a stationary problem involving a finite integral: the J. Bernoulli (1667-1748) curve of minimum descent time.

We look for the curve $y = f(x)$ such that the time t of descent is minimum. From simple mechanics,

$$t = \frac{1}{\sqrt{2g}} \int_a^b \frac{\sqrt{1+y'^2}}{\sqrt{a-y}} \, \mathrm{d}x \ .$$

The heights $\alpha = f(a)$ and $\beta = f(b)$ are given. This is an example of a variational problem. We look for $y(x)$ that makes t a minimum.

Euler (1707-1783) found a differential equation to solve the extreme of a integral between two point. Let's look for $y(x)$ such that

$$\int_{P_1}^{P_2} f(x, y(x), y'(x))\ \mathrm{d}x$$

presents a minimum (or in general, an extreme) between two fixed point P_1 and P_2.

Solution: Assume $y_0(x)$ is the solution. Take any close-by curve $y(x)$ between P_1 and P_2. It is assumed that $y(x) = y_0(x) + \epsilon\ \delta(x)$, where $\delta(x)$ is an arbitrary function that is zero at P_1 and P_2, and ϵ is a small parameter. The functions

$$J(\epsilon) = \int_{P_1}^{P_2} f(x,\ y_0(x) + \epsilon\ \delta(x),\ y_0'(x) + \epsilon\ \delta'(x))\ \mathrm{d}x$$

should represent a minimum (or in general, an extreme). Let us make $J'(\epsilon) = 0$. Since P_1 and P_2 are fixed, then

$$\frac{dJ(\epsilon)}{d\epsilon} = \int_{P_1}^{P_2} \Big[f_y \delta(x) + f_{y'} \delta'(x) \Big]\ \mathrm{d}x = 0\ .$$

Integrating by part the 2^{nd} term and knowing that $\delta(x) = 0$ at P_1 and P_2, we obtain:

$$\Big[f_y - (\frac{\mathrm{d}}{\mathrm{d}x} f_{y'}) \Big] \delta(x) = 0\ .$$

Since δx is any function, then

$$f_y - \frac{\mathrm{d}}{\mathrm{d}x}(f_{y'}) = 0\ ,$$

but

$$\frac{\mathrm{d}}{\mathrm{d}x} f_{y'}(x, y, y') = f_{y',x} + f_{y',y} y' + f_{y',y'} y''\ .$$

Thus, the Euler equation may be written as:

$$f_y - (f_{y',x} + f_{y',y} y' + f_{y',y'} y'') = 0\ . \qquad (0.7.1)$$

Let us generalize this to the case where the function may depend on several functions $y(x)$, say $y_1,\ y_2,\ \ldots$, etc, and the corresponding first derivatives $y_1',\ y_2',\ \ldots$, etc.

The function $J(\epsilon)$ will appear: $y_i \rightarrow y_i^0 {+}\epsilon\ \delta(x)$, $y_i' \rightarrow {y_i'}^0 {+}\epsilon\ \delta'(x)$, etc, with $i = 1, 2, \ldots$ and y_i^0 being the assumed solution.

$$J(\epsilon) = \int_{P_1}^{P_2} f(x,\ y_1^0{+}\epsilon\ \delta(x),\ y_2^0{+}\epsilon\ \delta(x),\ \ldots,\ {y_1'}^0{+}\epsilon\ \delta'(x), {y_2'}^0{+}\epsilon\ \delta'(x),\ \ldots)\ \mathrm{d}x\ .$$

Now,

$$\Big(\frac{\mathrm{d}J(\epsilon)}{\mathrm{d}\epsilon}\Big)_{\epsilon=0} = 0 \rightarrow$$

$$\int_{P_1}^{P_2} \Big[\Big(\frac{\partial f}{\partial y_1}\delta(x) + \frac{\partial f}{\partial y_1'}\delta'(x)\Big) + \Big(\frac{\partial f}{\partial y_2}\delta(x) + \frac{\partial f}{\partial y_2'}\delta'(x)\Big),\ \ldots\Big]\ \mathrm{d}x = 0\ .$$

Each parentheses is like the previous case, so this will mean that the condition implies that

$$f_{y_1} - (f_{y_1',x} + f_{y_1',y_1} y_1' + f_{y_1',y_1'} y_1'') = 0\ ,$$

$$f_{y_2} - (f_{y_2',x} + f_{y_2',y_2} y_2' + f_{y_2',y_2'} y_2'') = 0\ , \text{etc.},$$

or in general:

$$f_{y_i} - (f_{y_i',x} + f_{y_i',y_i} y_i' + f_{y_i',y_i'} y_i'') = 0\ . \qquad (0.7.2)$$

This formula may be applied to the motion of a particle in an n-dimensional space. In particular, it will be applied to a four-dimensional space to deal with the general relativity of Einstein. We will obtain the geodesic and also bring in an interesting relationship of the Christoffel symbol versus the derivatives of the fundamental metric tensor $g_{ij}(x)$ as a function of the four-dimensional coordinates.

Application of the Euler Differential Equation: Geodesic from the Variational Point of View

A geodesic line is by definition a line whose integral between two point P_1 and P_2 is stationary:

$$\delta \int_{P_1}^{P_2} \mathrm{d}s = 0$$

Assume a particle moving along the geodesic and consider the coordinates of the particle x^i as a function of one parameter λ, $x = x^i(\lambda)$. And instead of applying variational calculus to

$$\mathrm{d}^2 s = g_{ij}(x^k)\mathrm{d}x^i \mathrm{d}x^j \ ,$$

we will be doing it to

$$\mathrm{d}^2 w = g_{ij}(x^k(\lambda))\Big(\frac{\mathrm{d}x^i(\lambda)}{\mathrm{d}\lambda}\Big)\Big(\frac{\mathrm{d}x^j(\lambda)}{\mathrm{d}\lambda}\Big) \ .$$

In order to apply the Euler equation, we look at $\mathrm{d}^2 w$ as a function of λ where the functions $y_i(\lambda)$ become $x^i(\lambda)$ and its derivative $y_i'(\lambda)$ is $\dfrac{\mathrm{d}x^i(\lambda)}{\mathrm{d}\lambda}$, etc. The parameter λ plays the role of x in the previous derivation. The function $f(\lambda)$ is identified with $\mathrm{d}^2 w$. Thus, $f = f(\lambda, x^k(\lambda), \dfrac{\mathrm{d}x^k(\lambda)}{\mathrm{d}\lambda})$. From this,

$$\frac{\partial f}{\partial x^k} = g_{ij,k}\Big(\frac{\mathrm{d}x^i}{\mathrm{d}\lambda}\Big)\Big(\frac{\mathrm{d}x^j}{\mathrm{d}\lambda}\Big) \ ,$$

and

$$\frac{\mathrm{d}}{\mathrm{d}\lambda}\Big(\frac{\partial f}{\partial x^k}\Big) = \frac{\mathrm{d}}{\mathrm{d}\lambda}\Big(2g_{ij}\frac{\mathrm{d}x^i}{\mathrm{d}\lambda}\delta_{jk}\Big) \ .$$

The factor of 2 is due to the fact that i and j are exchangeable, since we assume a symmetric tensor g_{ij}.

Now,

$$\frac{\mathrm{d}}{\mathrm{d}\lambda}\Big(2g_{ij}\frac{\mathrm{d}x^i}{\mathrm{d}\lambda}\delta_{jk}\Big) = 2g_{ij,\sigma}\frac{\mathrm{d}x^\sigma}{\mathrm{d}\lambda}\frac{\mathrm{d}x^i}{\mathrm{d}\lambda}\delta_{jk} + 2g_{ij}\frac{\mathrm{d}^2 x^i}{\mathrm{d}\lambda^2}\delta_{jk} \ .$$

The Euler equation reads:

$$g_{ij,k}\frac{\mathrm{d}x^i}{\mathrm{d}\lambda}\frac{\mathrm{d}x^j}{\mathrm{d}\lambda} = 2\Big(g_{ij,\sigma}\frac{\mathrm{d}x^\sigma}{\mathrm{d}\lambda}\frac{\mathrm{d}x^i}{\mathrm{d}\lambda}\delta_{jk} + g_{ij}\frac{\mathrm{d}^2x^i}{\mathrm{d}\lambda^2}\delta_{jk}\Big)$$

By changing σ for j, and exchanging j for k in the first term of the right-hand side, we get:

$$g_{ij}\frac{\mathrm{d}^2x^i}{\mathrm{d}\lambda^2} + g_{ik,j}\frac{\mathrm{d}x^i}{\mathrm{d}\lambda}\frac{\mathrm{d}x^j}{\mathrm{d}\lambda} - \frac{1}{2}g_{ij,k}\frac{\mathrm{d}x^i}{\mathrm{d}\lambda}\frac{\mathrm{d}x^j}{\mathrm{d}\lambda} = 0$$

or

$$g_{ij}\frac{\mathrm{d}^2x^i}{\mathrm{d}\lambda^2} + (g_{ik,j} - \frac{1}{2}g_{ij,k})\frac{\mathrm{d}x^i}{\mathrm{d}\lambda}\frac{\mathrm{d}x^j}{\mathrm{d}\lambda} = 0\ .$$

Since j and i are interchangeable, $g_{ik,j} = g_{jk,i}$, and

$$g_{ik,j} = \frac{1}{2}(g_{ik,j} + g_{jk,i})\ .$$

Thus, the Euler equation becomes:

$$g_{ij}\frac{\mathrm{d}^2x^i}{\mathrm{d}\lambda^2} + \frac{1}{2}(g_{ik,j} + g_{jk,i} - g_{ij,k})\frac{\mathrm{d}x^i}{\mathrm{d}\lambda}\frac{\mathrm{d}x^j}{\mathrm{d}\lambda} = 0 \qquad (0.7.3)$$

The term in parentheses is the Chirstoffel symbol Γ_{ikj} ..

$$\Gamma_{ikj}{=}\frac{1}{2}[g_{ik,j} + g_{kj,i} - g_{ij,k}]\ .$$

Multiplying by g^{lk}, we have a new form of Christoffel symbol,

$$g^{lk}\Gamma_{ikj} = \Gamma^l_{ij}\ .$$

By direct analysis and since $g_{ij,k} = g_{ji,k}$, it is easy to see that Γ_{ikj} is symmetric in i and j. That is,

$$\Gamma_{ikj} = \Gamma_{jki}\ .$$

We must observe that Christoffel symbols are not tensors, because under a change of variables they do not change, as a tensor does (called in our previous lecture an *element*). They are not attached at a point in the space, but rather they indicate a bridge between two close points. However, we may lower or raise their indices, as we do with tensors, by using the metric tensor g_{ij}. Also, if a tensor varies in a coordinate system it does in any other; this is not so for Christoffel symbols.

Thus, the equation of motion of a point on a geodesic may be written as:

$$g_{ij}\frac{\mathrm{d}^2x^i}{\mathrm{d}s^2} + \Gamma_{ikj}\frac{\mathrm{d}x^i}{\mathrm{d}s}\frac{\mathrm{d}x^j}{\mathrm{d}s} = 0 \ ,$$

where we use $\mathrm{d}s$ instead of $\mathrm{d}\lambda$.

Changing i by l and j by k, in the first term it may be rewritten as:

$$g_{kl}\frac{\mathrm{d}^2x^l}{\mathrm{d}s^2} + \Gamma_{ikj}\frac{\mathrm{d}x^i}{\mathrm{d}s}\frac{\mathrm{d}x^j}{\mathrm{d}s} = 0 \ .$$

Multiplying by g^{kl},

$$\frac{\mathrm{d}^2x^l}{\mathrm{d}s^2} + g^{kl}\Gamma_{ikj}\frac{\mathrm{d}x^i}{\mathrm{d}s}\frac{\mathrm{d}x^j}{\mathrm{d}s} = 0 \ .$$

By using the new form of the Christoffel symbol, the equation of motion of a particle on a geodesic is:

$$\frac{\mathrm{d}^2x^l}{\mathrm{d}s^2} + \Gamma^l_{ij}\frac{\mathrm{d}x^i}{\mathrm{d}s}\frac{\mathrm{d}x^j}{\mathrm{d}s} = 0 \ . \tag{0.7.4}$$

where $\Gamma^l_{ij} = g^{kl}\Gamma_{ikj}$.

Geodesic and absolute derivatives

Having explained the equation of a geodesic using the Lagrange equation from variational calculus, next we will derive it from the concept of the absolute derivative of a vector developed previously.

Let us use for s, the length of the curve, time as the parameter. The velocity in contravariant form is then

$$v^i = \frac{\mathrm{d}q^i}{\mathrm{d}t}$$

Let us call ζ^i the contravariant component of the acceleration $\zeta^i = \dfrac{\mathrm{d}v^i}{\mathrm{d}t}$. The absolute derivative of this contravariant vector is

$$\mathrm{D}\zeta^i = \mathrm{d}\zeta^i + \zeta^j \Gamma^i_{kj} \mathrm{d}q^k$$

Dividing the differential by dt, we get

$$\frac{\mathrm{D}\zeta^i}{\mathrm{d}t} = \frac{\mathrm{d}\zeta^i}{\mathrm{d}t} + \zeta^j \Gamma^i_{kj} \frac{\mathrm{d}q^k}{\mathrm{d}t}$$

Now returning the original variables q^i and calling $\gamma^i = \dfrac{\mathrm{D}\zeta^i}{\mathrm{d}t}$,

$$\gamma^i = \frac{\mathrm{d}^2 q^i}{\mathrm{d}t^2} + \frac{\mathrm{d}q^j}{\mathrm{d}t} \Gamma^i_{kj} \frac{\mathrm{d}q^k}{\mathrm{d}t}$$

If this acceleration is equal to zero, then we have the equation for the geodesic; namely,

$$\frac{\mathrm{d}^2 q^i}{\mathrm{d}t^2} + \Gamma^i_{kj} \frac{\mathrm{d}q^j}{\mathrm{d}t} \frac{\mathrm{d}q^k}{\mathrm{d}t} = 0$$

The parameter t may be changed to s, and we have the same equation that was obtained before using variational calculus.

We may say that the velocity v has the same componenet at the point q and at its neighbor $q + \mathrm{d}q$. In other words, v moves parallel to itself, preserving the same contravariant components v^i and obviously tangent to the geodesic.

Also,

$$\frac{\mathrm{D}v}{\mathrm{d}s} = \frac{\mathrm{D}v}{\mathrm{d}q^k}\frac{\mathrm{d}q^k}{\mathrm{d}s}$$

where the parameter s represents the length on the geodesic. Since

$$\frac{\mathrm{D}v}{\mathrm{d}s} = 0$$

it means that

$$\frac{\mathrm{D}v}{\mathrm{d}q^k}\frac{\mathrm{d}q^k}{\mathrm{d}s} = \mathrm{D}_k\vec{v} \cdot \vec{r^k} = 0$$

where $\vec{r^k}$ is the unit vector along the tangent. In other words, $\dfrac{\mathrm{D}v}{\mathrm{d}q^k}$ is a vector perpendicular to the curve. As we will see later, Weyl generalizes the study of parallel displacement to the Riemann space and other paths which are not geodesic. And again, this approach is equivalent to the absolute derivative of the statement that the vector is null.

Geodesic: Summary

Let us summarize some of the main features of a geodesic line.

1. From the variational calculus point of view, the geodesic is the curve that represents the minimum distance between two points in the space. In a flat space it is just a straight line. Einstein associates gravitation with local or proper acceleration in empty space, but only if the space is curved may the geodesic be generated by a gravitational field.

   In a spherical surface the equator is a geodesic and the mobile will move with a velocity tangent to it.

2. A parallel transport of a vector is equivalent to having its absolute derivative equal zero. In this displacement, the vector may or may not preserve the values of its contravariant components. If it does, then the transport is done using coordinates that we will call *geodesic coordinates*. If it does not preserve the contravariant component, then the variation of the component from one point to another next to it is associated with the Christoffel symbol or connecting function.

3. The geodesic equation establishes that a particle in a Riemann Space will have absolute acceleration equal to zero. That is, the velocity on the path moves parallel to itself and its components are kept constant. Thus, if the velocity of the mobile is zero at the beginning of the path, it will continue to be zero in its ”parallel” transport. We say that the geodesic is a null one. This is a new way to express Newton’s First Law in Riemann Space: a particle with a constant velocity or at rest will continue to move at the same velocity or remain at rest. The particle is assumed to be in the gravitational field, but not under any other type of field.

4. The geodesic thus obtained has been derived from a variational principle, or from the tensor derivative introduced: one by Euler and the other by Christoffel. Neither approach makes use of any system of coordinates. This independence of the differential equation from the system of coordinates is much in agreement with the Principle of Relativity. Equations in physics derivable from a variational principal are called ”self-adjoint.”

D'Alembert Principle

D'Alembert (1771-1783) proposed to look at the dynamic of a point as a static case where the forces on the particle are of three kinds:

1. External forces F
2. Reaction force of the constraint R

   and
3. The inertial forces $(-m\dfrac{d^2x}{dt^2} = I)$.

Thus, if $\vec{I}$ is the vector representing the inertial forces, and $\mathcal{F}$ represents the external forces $F + R$, the reaction from the constraint, then m is in perfect static equilibrium. When moving m by some arbitrary displacement (which we call virtual displacement), the work of $\vec{I} + \mathcal{F}$ is 0. This virtual displacement is done at a given time. They may or may not coincide with the real displacement. For instance, if the point moves on a curve c, the virtual displacement is $A - B$. But if the constraint depends on time and c moves to c' in Δt, the real displacement is $A - C$, which is different from $A - B$.

As we refer to virtual displacement we are considering small differential displacement δx, δy, δz, or in generalized coordinates, δq.

If the displacement is compatible with the constraint, and this is frictionless, the virtual work due to the external forces F and the inertial forces is null.

For instance, in the x direction

$$\Big(F_x - m\frac{d^2x}{dt^2}\Big)\delta x = 0 \ .$$

Between the external force and the inertial force there are, from the mathematical point of view, great differences: F may be derived from a potential $V(x, y, z)$, so $F\delta x = -\Delta V$; however, the inertial forces are not derivable from a function and we may have to treat each point of the system individually.

In order to make the inertial force derivable from a function, Euler (1707-1783), Lagrange (1736-1813), and finally Hamilton (1805-1865), transformed the approach of D'Alembert into a variational problem. We have seen that Jakob Bernoulli (1657-1705) and his brother (1667-1748) tried the problem of finding the path that will take the least time for a particle which travels from A to B under the action of gravity.

This problem implies the variation of an integral, and so falls into what we call variational calculus. Lagrange changed this problem of minimizing an integral into a differential equation. As we saw, the differential equation that solves the function $f(x, y(x), y'(x))$ such that $\int_A^B f(x, y(x), y'(x))dx$ is a minimum (or maximum) is

$$\frac{\partial f}{\partial y} - \frac{\mathrm{d}}{\mathrm{d}x}\frac{\partial f}{\partial y'} = 0 \ ,$$

where

$$\frac{\mathrm{d}}{\mathrm{d}x}\frac{\partial f}{\partial y'} = \left(\frac{\partial f}{\partial y'}\right)\frac{\partial y'}{\partial x} + \left(\frac{\partial f}{\partial y'}\right)\frac{\partial y}{\partial x} + \frac{\partial f}{\partial y'} \ ,$$

since

$$\frac{\partial f}{\partial y'} = g(x, y(x), y'(x)) \ .$$

In order to bring in the variational aspect of the dynamic equation, we integrate between two times t_1 and t_2, and the virtual differential works as expressed by D'Alembert:

$$\tau = \int_{t_1}^{t_2} \Big[F - \frac{\mathrm{d}^2}{\mathrm{d}t^2}(mx)\Big]\delta s \ \mathrm{d}t \ .$$

If F is derivable from a potential V $\delta V = -F\delta x$, then the first term of the integral is:

$$\int_{t_1}^{t_2} \delta V \ \mathrm{d}t = \delta\int_{t_1}^{t_2} (-V) \ \mathrm{d}t \ .$$

The second term may be integrated by part:

$$\int_{t_1}^{t_2} \frac{\mathrm{d}}{\mathrm{d}t}(m\dot{x})\delta x \ \mathrm{d}t = \Big[\delta x(m\dot{x})\Big]_{t_1}^{t_2} - \int_{t_1}^{t_2} m\dot{x}\frac{\mathrm{d}}{\mathrm{d}t}\delta x \ \mathrm{d}t \ .$$

Taking $\delta x = 0$ at t_1 and t_2, the first term is zero, and the second, noting that the derivative of the variation δx is equal to the variation of the derivative, is: $\frac{\mathrm{d}}{\mathrm{d}t}(\delta x) = \delta(\frac{\mathrm{d}}{\mathrm{d}t}x) = \delta\dot{x}$.

$$-\int_{t_1}^{t_2} m\dot{x}(\delta\dot{x})\ dt = -\int_{t_1}^{t_2} \delta\Big(\frac{m\dot{x}^2}{2}\Big)\ dt = -\delta\int_{t_1}^{t_2} T(t)\ \mathrm{d}t\ ,$$

where T is the kinetic energy. So:

$$\tau = \delta\int_{t_1}^{t_2} (-V + T)\ \mathrm{d}t$$

We indicate the derivative of any variable with respect to t by an upper dot $(\cdot)$. By calling $\mathcal{L} = (T - V)$ (the excess of kinetic energy over potential energy) the Lagrange function, the problem is reduced to finding the $\mathcal{L}$ that make $\int_{t_1}^{t_2} \mathcal{L}\ \mathrm{d}t$ a minimum or stationary between two instants t_1 and t_2. τ is called the action integral. Using the Euler differential equations, we find the function $x(t)$ from

$$\frac{\partial\mathcal{L}}{\partial x} - \frac{\mathrm{d}}{\mathrm{d}t}\frac{\partial\mathcal{L}}{\partial\dot{x}} = 0\ .$$

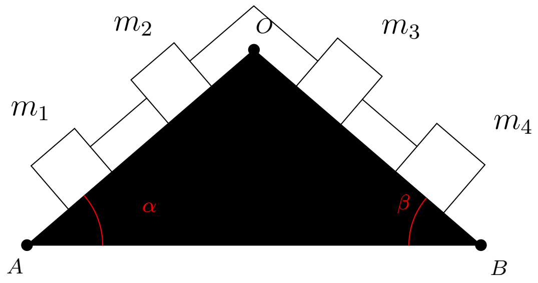
If we are working with generalized coordinates q' , $q^2 \ldots q^n$, then

$$\frac{\mathrm{d}}{\mathrm{d}t}\frac{\partial\mathcal{L}}{\partial\dot{q}^i} - \frac{\partial\mathcal{L}}{\partial q^i} = 0\ , \qquad (0.7.5)$$

and these are n equations $(i = 1 \ldots n)$.

The equation is called the Euler-Lagrange equation.

Problems Using the D'Alembert Principle and the Euler-Lagrange Equation



1.

Four masses slide down on the surface of the triangle without friction. They are connected by a massless thread. Assume they slide positively from B to A through O. The coordinates will be called q, which is the distance from O to the point masses. The virtual work will be due to a variation δq which in this case is similar to a real one (a slide down the surface).

The work of the inertial force is:

$$m_1\frac{\partial^2 q}{\partial t^2}\delta q + m_2\frac{\partial^2 q}{\partial t^2}\delta q - m_3\frac{\partial^2 q}{\partial t^2}\delta q - m_4\frac{\partial^2 q}{\partial t^2}\delta q$$

while the work of the external force (gravity) is

$$m_1 g\sin\alpha\delta q + m_2 g\sin\alpha\delta q - m_3\sin\beta\delta q - m_4\sin\beta\delta q$$

We assume that the motion is direction downwards for m_1 and m_2, as are the forces acting on these objects. For m_3 and m_4, the force is opposite in direction to the displacement, so according to the D'Alembert Principal, the sum of these two works is zero. That is,

$$g\sin\alpha(m_1+m_2) - g\sin\beta(m_3+m_4) = (m_1+m_2+m_3+m_4)\frac{\partial^2 q}{\partial t^2}$$

Only if the left side of the equation is null will the motion be constant with uniform acceleration.

Now let us solve the same problem using the Euler-Lagrange Equation, $\mathcal{L} = T - V$, where $T = \frac{1}{2}\sum_{i=1}^{4} m_i \dot{q}_i^2$. Assuming zero potential at the base of the triangle ABO and taking h as the height,

$$V_1 = (h - q_1 \sin\alpha) m_1 g \quad , \quad V_2 = (h - q_2 \sin\alpha) m_2 g \quad , \quad V_3 = (h + q_3 \sin\beta) m_3 g$$
$$V_4 = (h + q_4 \sin\alpha) m_4 g$$

[redacted]

Now

$$\frac{\partial V_1}{\partial q_1} = -(\sin\alpha) m_1 g \quad , \quad \frac{\partial V_2}{\partial q_2} = -(\sin\alpha) m_2 g$$
$$\frac{\partial V_3}{\partial q_3} = (\sin\beta) m_3 g \quad , \quad \frac{\partial V_4}{\partial q_4} = (\sin\beta) m_4 g$$

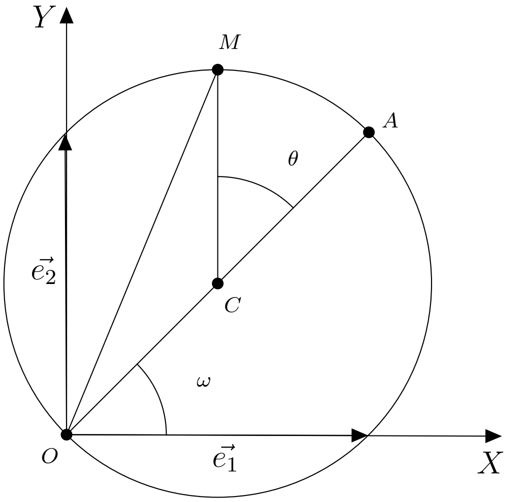
Since $\mathrm{d}q_1 = \mathrm{d}q_2 = -\mathrm{d}q_3 = -\mathrm{d}q_4$,

$$\frac{\partial \mathcal{L}}{\partial q} = -\frac{\partial V}{\partial q} = g[(m_1 + m_2)\sin\alpha - (m_3 + m_4)\sin\beta]$$

and
$$\frac{\mathrm{d}}{\mathrm{d}t}\frac{\partial \mathcal{L}}{\partial \dot{q}} = \frac{\mathrm{d}}{\mathrm{d}t} T = \sum_i m_i \frac{\mathrm{d}}{\mathrm{d}t}\dot{q} = M\ddot{q}$$

where $M = \sum m_i$. Thus, the Euler-Lagrange Equation $\frac{\partial \mathcal{L}}{\partial q} - \frac{\mathrm{d}}{\mathrm{d}t}\frac{\partial \mathcal{L}}{\partial \dot{q}} = 0$ gives

$$M\ddot{q} = g[(m_1 + m_2)\sin\alpha - (m_3 + m_4)\sin\beta]$$



2.

The point M moves on a horizontal circumference, varying the angle θ. The circumference rotates at the same time around a point O fixed to it with angular velocity ω. There are two variables: $q^1 = \theta$ and $q^2 = \omega t$. We assume that ω is constant. So,

$$\dot{q}^1 = \dot{\theta} \quad \text{and} \quad \dot{q}^2 = \omega$$

Since M moves on a horizontal plane, there is no potential variation. Now, $T = \frac{1}{2}m(\dot{x}^2 + \dot{y}^2)$ where m is the mass of M. Also,

$$x = \overrightarrow{OM} \cdot \vec{e_1} = (\overrightarrow{OC} + \overrightarrow{CM}) \cdot \vec{e_1} = R\cos(\omega t) + R\cos(\theta + \omega t)$$

where R is the radius of the circle, and

$$y = \overrightarrow{OM} \cdot \vec{e_2} = (\overrightarrow{OC} + \overrightarrow{CM}) \cdot \vec{e_2} = R\sin(\omega t) + R\sin(\omega t + \theta)$$

Thus,

$$\begin{aligned}\dot{x} &= -R\omega\sin(\omega t) - R\sin(\omega t+\theta)(\omega+\dot{\theta})\\ \dot{y} &= R\omega\cos(\omega t) + R\cos(\omega t+\theta)(\omega+\dot{\theta})\end{aligned}$$

$$T = \frac{1}{2}m(R^2\omega^2 + R^2(\omega+\dot{\theta})^2 + 2R^2\omega(\omega+\dot{\theta})[\sin(\omega t)\sin(\omega t+\theta)+\cos(\omega t)\cos(\omega t+\theta)])$$

$$T = \frac{1}{2}mR^2[\omega^2 + (\omega+\dot{\theta})^2 + 2\omega(\omega+\dot{\theta})\cos\theta]$$

[redacted]

Now, since v is a constant, the Euler-Lagrange Equation is

$$\frac{\partial\mathcal{L}}{\partial q} - \frac{\mathrm{d}}{\mathrm{d}t}\frac{\partial\mathcal{L}}{\partial\dot{q}} = \frac{\partial T}{\partial q} - \frac{\mathrm{d}}{\mathrm{d}t}\frac{\partial T}{\partial\dot{q}} = 0$$

There are two equations, one for $q=\theta$ and the other for $q=\omega t$:

$$\begin{aligned}\frac{\partial T}{\partial\theta} &= \frac{1}{2}mR^2(-2\omega(\omega+\dot{\theta})\sin\theta) = -mR^2\omega(\omega+\dot{\theta})\sin\theta\\ \frac{\mathrm{d}}{\mathrm{d}t}\frac{\partial T}{\partial\dot{\theta}} &= mR^2\frac{\mathrm{d}}{\mathrm{d}t}[(\omega+\dot{\theta})+\omega\cos\theta] = mR^2(\ddot{\theta}-\omega\dot{\theta}\sin\theta)\\ \frac{\partial T}{\partial\theta} - \frac{\mathrm{d}}{\mathrm{d}t}\frac{\partial T}{\partial\dot{\theta}} &= -\omega^2\sin\theta - \omega\dot{\theta}\sin\theta - \ddot{\theta} + \omega\dot{\theta}\sin\theta = 0\end{aligned}$$

or

$$\ddot{\theta} = -\omega^2\sin\theta$$

This is the equation of a pendulum as seen from the center of the circle.

Euler–Lagrange–Poisson–Jacobi–Hamilton–Canonical Equations

The Euler-Lagrange equation is:

$$\frac{\mathrm{d}}{\mathrm{d}t}\frac{\partial T}{\partial \dot{q_\nu}} - \frac{\partial T}{\partial q_\nu} = Q_\nu \ ,$$

where Q_ν is the force and $\dot{q_\nu} = \dfrac{\mathrm{d}q_\nu}{\mathrm{d}t}$.

Assume that q_1, q_2, q_ν, q_k are the degrees of freedom of the system of n points where there are h equations of constraint such that:

$$k = 3n - h \ .$$

In the following we simplify our equations by eliminating the subscript ν and using such q as a general variable. With the same criteria we will use a comma to indicate a partial derivative like $\dfrac{\partial T}{\partial q}$ as T, q.

The kinetic energy T is given by $T = a^{\nu\eta}\dot{q_\nu}\dot{q_\eta}$, where the matrix $a^{\nu\eta}$ is a function of the coordinates q_1, q_2, $\ldots$. Assuming $T = T(\dot{q}, q)$, its variation is given by:

$$\delta T = T, \dot{q}\ \delta\dot{q} + T, q\ \delta q \ .$$

By defining p as:

$$p = \frac{\partial T}{\partial \dot{q}} \ ,$$

$$\delta T = p\ \delta\dot{q} + T, q\ \delta q \ .$$

By summing and subtracting $\dot{q}\ \delta p$, then:

$$\delta(\dot{q}p - T) = \delta K = \dot{q}\ \delta p - T, q\ \delta q \ ,$$

where we have defined K as:

$$K = \dot{q}p - T$$

If K is considered a function of q and p only:

$$\frac{\partial K}{\partial q} = -\frac{\partial T}{\partial q}$$

and

$$\frac{\partial K}{\partial p} = \dot{q} \ .$$

From the Euler-Lagrange equation written above and the definition of K,

$$\frac{dp}{dt} + \frac{\partial K}{\partial q} = Q$$

or

$$\dot{p} + \frac{\partial K}{\partial q} = Q \ .$$

For $Q = 0$,

$$\dot{p} = -\frac{\partial K}{\partial q} \ .$$

If Q is derivable from a potential V such that

$$Q = -\frac{\partial V}{\partial q}$$

where $V = V(q,t)$, then by defining $\mathcal{H} = (K + V)$,

$$\frac{\partial \mathcal{H}}{\partial p} = \dot{q}$$

and

$$\frac{\partial \mathcal{H}}{\partial q} = \frac{\partial K}{\partial q} + \frac{\partial V}{\partial q} = -\frac{\partial}{\partial q}T - Q \ .$$

By using the Euler-Lagrange equation and the definition of p, $p = \dfrac{\partial T}{\partial q}$,

$$\frac{\partial \mathcal{H}}{\partial q} = -\dot{p} \ .$$

These two equations, where p and q play symmetric roles, are called the Canonical Equations, or Hamilton Equations:

$$\frac{\partial \mathcal{H}}{\partial p} = \dot{q} \quad \text{and} \quad \frac{\partial \mathcal{H}}{\partial q} = -\dot{p} \ .$$

The variables are $q_1, q_2, \ldots, q_k$ and $p_1, p_2, \ldots, p_k$.

Lagrange–Poisson–Jacobi–Maupertuis

After the Lagrangian $\mathcal{L}$, other interesting functions were created whose first derivatives will give rise to the equation of motion. They assume different sets of variables. The generalized variables $q_1, q_2, \ldots, t$ that we represent by (q, t), were used by Lagrange in his function $\mathcal{L}(q, t)$ and Jacobi in his function $S(q, t)$.

As we have seen, $\mathcal{L}$ is $(T - V)$, the excess of kinetic energy over the potential, and it holds the equation

$$\frac{\mathrm{d}\partial\mathcal{L}}{\mathrm{d}t\partial\dot{q}} - \frac{\partial\mathcal{L}}{\partial q} = 0 \; .$$

The function S of Jacobi is given by

$$S(q, t) = -\int_0^t \mathcal{L}(q, t) \; \mathrm{d}t \; .$$

If $\mathcal{L} = (T - V)$ is written as $\mathcal{L} = -(T + V - p\dot{q})$, where $p\dot{q}$ indicates $p_i\dot{q}_i = 2T$ (repeated indices implies summation), then:

$$S(q, t) = \int_0^t [(T + V) - p\dot{q}] \; \mathrm{d}t$$

By saying that $\mathcal{H} = T + V$, we then rewrite the previous equation as

$$S(q, t) = \int_0^t \mathcal{H} \; dt - \int_{q_0}^q p \; \mathrm{d}q$$

which generates two first order fundamental derivatives:

$$\mathcal{H} = \frac{\partial S}{\partial t} \quad \text{and} \quad \frac{\partial S}{\partial q} = -p$$

These derivatives resemble the Schrodinger recipe for quantization:

$$\mathcal{H} \to \frac{\partial}{\partial t} \quad \text{and} \quad p \to -\frac{\partial}{\partial q}$$

The Jacobi Equation that will give $q = q(t)$ is

$$\frac{\partial S}{\partial t} = \mathcal{H}(q, p, t) \quad \text{where} \quad p = -\frac{\partial S}{\partial q}$$

Thus, the Jacobi Equation reads:

$$\mathcal{H}(q,-\frac{\partial S}{\partial q},t)-\frac{\partial S}{\partial t}=0 \tag{0.7.6}$$

where $S(q,t)$ will be equivalent to solve the particle motion.

To show this, let us see that $S(q,t)$ generates the Canonical Equations. In fact, a solution of the Jacobi equation will give $S = S(q,t,a)$ where a is a constant of integration (one for each coordinate of q). By deriving S with respect to a we generate a new constant b such that

$$\frac{\partial S}{\partial a}=b \tag{0.7.7}$$

In $S(q,t,a)$ we have $q = q(t)$ for each a. Then from

$$\frac{\partial S}{\partial q}=-p \tag{0.7.8}$$

we get $p(q,t)$.

Let us differentiate (0.7.7) with respect to t and a. From (0.7.7), since b is a constant, we have

$$\frac{\partial^2 S}{\partial t\partial a}+\frac{\partial^2 S}{\partial a\partial q}\dot{q}=0$$

By differentiating (0.7.6) with respect to a, which is implicit in

$$\frac{\partial S}{\partial t} \quad \text{and} \quad \frac{\partial S}{\partial q}\ ,$$

$$-\frac{\partial^2 S}{\partial a\partial t}-\frac{\partial \mathcal{H}}{\partial(\frac{\partial S}{\partial q})}\frac{\partial^2 S}{\partial a\partial q}=0$$

Equating these two equations, we see that

$$\dot{q}=\frac{\partial \mathcal{H}}{\partial(-\frac{\partial S}{\partial q})}$$

But we saw that

$$-\frac{\partial S}{\partial q} = p$$

so

$$\dot{q} = \frac{\partial H}{\partial p}$$

which is one of the Canonical Equations.

To obtain the second Canonical Equation, we differentiate (0.7.8) with respect to t and (0.7.6) with respect to q and compare:

$$\frac{\partial^2 S}{\partial q \partial t} + \frac{\partial^2 S}{\partial^2 q}\dot{q} = -\dot{p} \tag{0.7.9}$$

$$-\frac{\partial^2 S}{\partial t \partial q} - \frac{\partial \mathcal{H}}{\partial(\frac{\partial S}{\partial q})} \frac{\partial(\frac{\partial S}{\partial q})}{\partial^2 q} + \frac{\partial \mathcal{H}}{\partial q} = 0$$

Equating, after a change of sign in the last equation, the second term gives rise to the previous Canonical Equation and the third gives

$$\frac{\partial \mathcal{H}}{\partial q} = -\dot{p}$$

which is the second Canonical Equation.

In this form, we have shown that the Jacobi Equation creates $S(q,t)$ that solves the dynamical problem in question.

Maxwell Equations and Their Invariance Under Gauge Change[redacted]

The Maxwell Equations are based on the following four fundamental equations:

1. Faraday:
$$\nabla \times E + \frac{\partial B}{\partial t} = 0$$
2. Gauss:
$$\nabla \cdot E = \frac{\rho}{\epsilon_0}$$
3. Ampere:
$$\nabla \times B = J$$
4. No free magnetic pole:
$$\nabla \cdot B = 0$$

where
$$H = \frac{B}{\mu_0} \quad , \quad D = \epsilon_0 E \quad , \quad \sqrt{\mu_0 \epsilon_0} = \frac{1}{c}$$

Since $\nabla \cdot \nabla \times = 0$, Ampere would say that $\nabla \cdot J = 0$. But the continuity of charge, $\nabla \cdot J + \frac{\partial E}{\partial t} = 0$, must be fulfilled. So Maxwell changed Ampere's equation to the Ampere-Maxwell Equation:

$$\nabla \times B - \frac{\partial E}{c^2 \partial t} = \mu_0 J$$

Now, if we apply $\nabla \cdot$ to both sides we generate the continuity equation.

So the Maxwell Equations are then:

$$\nabla \times E = -\frac{\partial B}{\partial t} \tag{0.7.10}$$

$$\nabla \cdot E = \frac{\rho}{\epsilon_0} \tag{0.7.11}$$

$$\nabla \times B = \mu_0 J + \frac{\partial E}{c^2 \partial t} \tag{0.7.12}$$

$$\nabla \cdot B = 0 \tag{0.7.13}$$

These four equations can be simplified by introducing the EM Potentials: $\vec{A}$ and Φ:

$$B = \nabla \times \vec{A} \tag{0.7.14}$$

$$E = -\nabla\Phi - \frac{\partial A}{\partial t} \tag{0.7.15}$$

From (0.7.13) and (0.9.51),

$$\nabla \cdot \nabla \times A = 0$$

From (0.7.15) and (0.7.11),

$$\nabla \cdot E = -\nabla \cdot (\nabla\Phi + \frac{\partial A}{\partial t}) = -\nabla^2\Phi - \frac{\partial \nabla}{\partial t} \cdot A = \frac{\rho}{\epsilon}$$

or

$$\nabla^2\Phi + \frac{\partial}{\partial t}(\nabla \cdot A) = -\frac{\rho}{\epsilon} \tag{0.7.16}$$

Substituting (0.9.51) and (0.7.15) in (0.7.10),

$$\nabla \times (-\nabla\Phi - \frac{\partial A}{\partial t}) = -\frac{\partial}{\partial t}\nabla \times A \rightarrow \text{all cancel since } \nabla \times \nabla\Phi = 0$$

Substituting (0.9.51) and (0.7.15) in (0.7.12),

$$\nabla \times \nabla \times A = \mu_0 J + \frac{\partial}{c^2 \partial t}(-\nabla\Phi - \frac{\partial A}{\partial t}) \tag{0.7.17}$$

Observe that

$$\nabla \times \nabla \times A = \nabla(\nabla \cdot A) - \nabla^2 A$$

or

$$\nabla^2 A - \frac{\partial^2 A}{c^2 \partial t^2} - \nabla(\nabla \cdot A + \frac{\partial}{c^2 \partial t}\Phi) = -\mu_0 J \; . \tag{0.7.18}$$

(0.7.16) and (0.7.18) ar coupled equations for A and Φ.

A and Φ may be changed without varying Maxwell's Equations (0.7.16) and (0.7.18):

$$A \to A' = A + \nabla\lambda$$

and

$$\Phi \to \Phi' = \Phi - \frac{\partial\lambda}{\partial t}$$

where λ is an arbitrary differential function. These gauge transformations leave (0.9.51) and (0.7.15) unchanged, and thus B and E are invariant. In four-dimensional tensor notation, these transformations may written as $A^\mu \to A^\mu - \partial^\mu\lambda$.

Lorentz fixed the gauge by requiring that

$$\nabla \cdot A + \frac{\partial\Phi}{c^2\partial t} = 0$$

or in four-dimensional tensor notation, $\partial_\mu A^\mu = 0$. This choice decouples (0.7.16) and (0.7.18) to the following two Poisson Equations in four-dimensional space:

$$\Box\Phi = -\rho \quad \text{and} \quad \Box\vec{A} = -J$$

or

$$\nabla^2\Phi - \frac{\partial^2\Phi}{c^2\partial t^2} = -\rho \quad \text{and} \quad \nabla^2 A - \frac{\partial^2 A}{c^2\partial t} = -\mu_0 J\ .$$

This is the so-called Coulomb Gauge, given by the condition that $\nabla \cdot A = 0$.

Tensor Form of the Maxwell Equations

From the potential equations ((0.9.51) and (0.7.15)), an antisymmetric tensor $F_{\mu\nu}$ is generated: assuming the coordinates $x_\mu : (x^0 = t, x^1 = x, x^2 = y, x^3 = z$ and $\partial_0 = \partial_t, \partial_1 = \partial_x, \partial_2 = \partial_y, \partial_3 = \partial_z)$,

$$F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu$$

where μ and ν run over the indices 0, 1, 2, 3 that represent the four dimensional space with the tensor potential A_μ: $A_0 = \Phi$ and $A_1 = A_x$, $A_2 = A_y$, and $A_3 = A_z$. Thus, $F_{\mu\nu}$ is given according to (0.9.51) and (0.7.15) by the covariant tensor $F_{\mu\nu}$, which may be expressed in its contravariant components by means of the metric tensor

$$g_{ij} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix}$$

As was shown in Lecture 2,

$$g^{ij} g_{ij} = 1$$

and thus

$$F^{ij} = F_{lm} g^{il} g^{jm} \ .$$

It results as

$$F^{ij} = \begin{pmatrix} 0 & -E^1 & -E^2 & -E^3 \\ E^1 & 0 & B^3 & -B^2 \\ E^2 & -B^3 & 0 & B^1 \\ E^3 & B^2 & -B^1 & 0 \end{pmatrix}$$

Weyl deduced this tensor from the connecting function Φ_μ, such that $\mathrm{d}l = \Phi_\mu \mathrm{d}q^\mu$, which has a non-zero rotational. So the circulation of the scalar l around a closed curve is non-zero. Thus,

$$F_{\mu\nu} = (\Phi_{\mu,\nu} - \Phi_{\nu,\mu})$$

has the following property:

$$\partial_\rho F_{\mu\nu} + \partial_\nu F_{\rho\mu} + \partial_\mu F_{\nu\rho} = 0 \ .$$

From the two-index tensor $F^{\mu\nu}$, we may generate a one-index tensor (or a vector) as

$$\partial_\nu F^{\mu\nu} = 4\pi J^\mu$$

where J^μ is the four-dimensional current $J^\mu = (\rho, \vec{j})$. Thus, by identifying Φ_μ with A_μ we regain the Maxwell Equations.

As we will see, Weyl makes a connection between matter (not gravitation as said in an earlier work) and the EM fields which move the particle out of the gravitational geodesic.

Let us assume a potential A_μ and shift it via the gradient of a function Ψ:

$$A_\mu(x) \rightarrow A'_\mu(x) = A_\mu(x) + \partial_\mu \Psi(x) \ .$$

Since the relevant property of A, that generates the neutral fields E and B, is its circulation, the gradient does not make any contribution and leaves the fields E and B invariant.

This transformation of the vector potential A_μ, which is a non-observable quantity, may be seen as a gauge transformation. (Lorentz, [redacted] in his famous work on EM, assumes that $\partial_\mu \Psi(x) = 0$). That is, a change in the connecting function Φ_μ, or a gauge transformation, does not change the Maxwell Equations. In other words, the tensor $F_{\mu\nu}$ has the same form as functions of the field E and B.

The tensor $F^{\mu\nu}$ gives rise to some invariants:

$$\text{Invariant (1)} : F_{ik}F^{ik}$$

$$\text{Invariant (2)} : e^{iklm}F_{ik}F_{lm}$$

where e^{iklm} is the four-rank antisymmetric unit tensor. Invariant (1) gives rise to: $2(B^2 - E^2)$ and Invariant (2) to: $(\vec{E} \cdot \vec{B})$.

Note: Calling $\vec{\Omega}$ the four-potential vector (Φ_0, A_x, A_y, A_z), the Maxwell Equations may be reduced to two equations:

$$\Box \vec{\Omega} = -\mu_0 \Gamma$$

and

$$\nabla \cdot \vec{\Omega} = 0$$

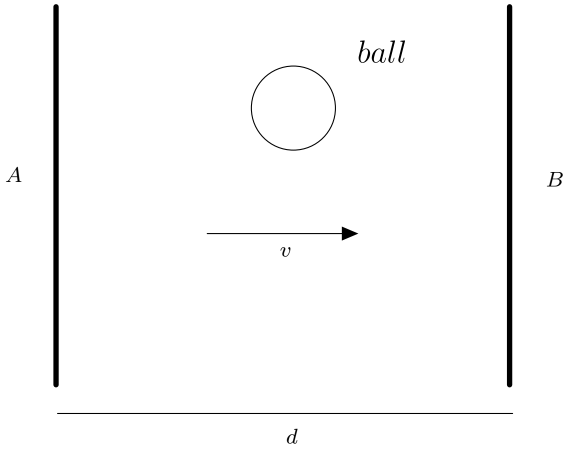
where

$$\Gamma = c\rho, J_x, J_y, J_z$$

Review of Relativity

1. **Concept of Simultaneity: Transmission of Information Between Clocks**

   Let us carry out the following experiment: A and B are two walls and a ball goes from A to B with velocity v.



At time t_0, the ball left A, and at time t_1, according to B, reached B. As the ball reaches B there is a sound that reflect back to A. The sound travels with velocity c and reaches A at time t_2 (according to A).

Time t_2 must depend on the velocity as

$$(t_2 - t_0) = \frac{d}{c} + \frac{d}{v} = d(\frac{1}{c} + \frac{1}{v})$$

If $c >> v$, then

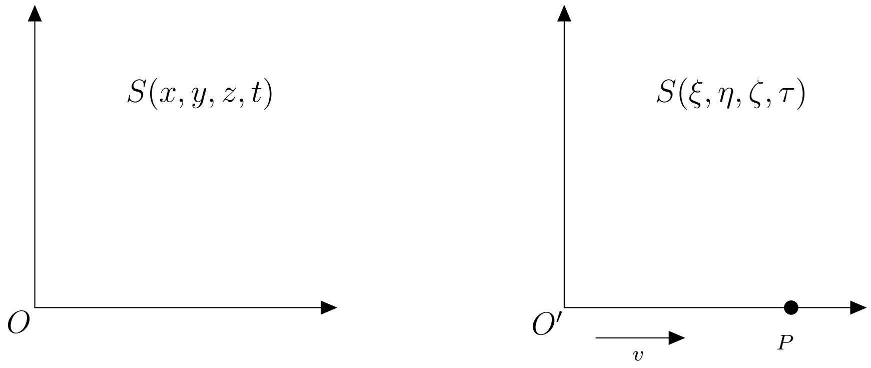
$$t_2 - t_0 \approx \frac{d}{v}$$

But if v is comparable with c, then the problem of time will depend on c in the ration $\frac{v}{c}$. In other words, the sound is the element to determine the simultaneity of an event observed by

A and also by B. If B moves away from A at the same time that the experiment is carried out, the sound speed for A is different from the one measured by an observer at B.

Einstein took light as the element for the correlation of time, and postulated that this velocity is the same for A as it is for B: a radical new concept.

2. **Simple Derivation of the Lorentz-Einstein Transformation**



The system S' moves with respect to S with a constant velocity v. Let us take a point P in the space of S' with coordinates $(\xi,0,0,\tau)$ and assume that S finds P to have coordinates $(x,0,0,t)$. Our next assumptions are

(a) The transformation from the S coordinates to the S' coordinates is a linear one.

(b) At the limit of low velocity v, the transformation is equal to the Galilean one.

(c) The space is isotropic.

(d) The velocity of light is equal for both observers, even when the light was emitted when O and O' were together and the light was moving away from O'.

Thus, from (a), assuming a, b, α, and β are constants,

$$\xi = ax + bt \quad \text{and} \quad \tau = \alpha x + \beta t$$

From (b), $b = -av$, or

$$\xi = a(x - vt) \tag{0.7.19}$$

From (c), we may exchange x for ξ, t for τ, and v for $-v$:

$$x = a(\xi + v\tau) \tag{0.7.20}$$

From (d), we may write

$$x = ct \quad \text{and} \quad \xi = c\tau \tag{0.7.21}$$

From (0.7.19), (0.7.20), and (0.7.21), we may determine the value of a:

$$c\tau = \xi = a(ct - vt) = at(c - v) \quad \text{and} \quad ct = x = a(c\tau + v\tau) = a\tau(c + v)$$ [redacted]

By multiplying these two equations, we obtain

$$c^2(t\tau) = a^2(\tau t)(c - v)(c + v)$$

or

$$a^2 = \frac{c^2}{c^2 - v^2} = \frac{1}{1 - \dfrac{v^2}{c^2}} = \frac{1}{1 - \beta^2}$$

where $\beta = \dfrac{v}{c}$. (In the usual notation, a is called γ).

By using (0.7.19) and (0.7.20) again, we obtain

$$\xi = \sqrt{\frac{1}{1 - \beta^2}}(x - vt) \tag{0.7.22}$$

and

$$x = \sqrt{\frac{1}{1 - \beta^2}}(\xi + v\tau) \tag{0.7.23}$$

From these equations, we solve for τ by substituting ξ into the second equation,

$$x = \gamma(\gamma(x - vt) + v\tau)$$

So

$$\tau = \frac{1}{v}\Big[\frac{x}{\gamma} - \gamma x + \gamma vt\Big] = \frac{x}{v}(\frac{1}{\gamma} - \gamma) + \gamma t$$

or

$$\tau = \gamma\Big(t - \frac{vx}{c^2}\Big) \tag{0.7.24}$$

By changing t for τ, x for ξ, and v for $-v$,

$$t = \gamma\Big(\tau + \frac{v\xi}{c^2}\Big) \tag{0.7.25}$$

Equations (0.7.22), (0.7.23), (0.7.24) and (0.7.25) express the Lorentz-Einstein Transformation. In matrix form, we write

$$\begin{pmatrix} x \\ t \end{pmatrix} = \gamma \begin{pmatrix} 1 & v \\ \dfrac{v}{c^2} & 1 \end{pmatrix} \begin{pmatrix} \xi \\ \tau \end{pmatrix}$$

or

$$\begin{pmatrix} \xi \\ \tau \end{pmatrix} = \gamma \begin{pmatrix} 1 & -v \\ -\dfrac{v}{c^2} & 1 \end{pmatrix} \begin{pmatrix} x \\ t \end{pmatrix}$$

The product of the two matrices, is, of course, equal to the unit matrix.

3. Variation of Length and Time from the S' Frame to the S Frame

From small increments of the coordinates, we may write the following equation:

$$\Delta\xi = \gamma(\Delta x - v\Delta t)$$

When the variation of the length is taken at a given time, $\Delta t = 0$ and

$$\Delta x = \frac{\Delta\xi}{\gamma}$$

Since $\gamma > 1$, $\Delta x < \Delta\xi$.

We say that, as $\Delta\xi$ moves away, its length looks contracted from the point of view of S. Similarly, as

$$\Delta t = \gamma\Big(\Delta\tau + \frac{v}{c^2}\Delta\xi\Big)$$

at one point in the space, the clock at S' moves slowly. That is, as $\Delta\xi = 0$,

$$\Delta t = \gamma\Delta\tau$$

and so $\Delta t > \Delta\tau$. Thus, there is a time dilation from $\Delta\tau \equiv \Delta t$ when the two clocks were at rest to a Δt larger than $\Delta\tau$ as O' moves away. In other words, the time interval measured by the moving clock will be small compared with that of the standard clock at rest, in spite of the fact that they were synchronized when O and O' were at rest. The two phenomena observed at S, time dilation and length contraction, are just velocity effects which become more significant when v approaches c (the velocity of light) in magnitude.

4. **The Differential Length of the Four-Dimensional Space as an Invariant**

Let us write $\mathrm{d}s = \mathrm{d}\tau c$, where τ is the "proper time" and c the speed of light. As we have shown,

$$\tau = \gamma\Big(t - \frac{v}{c^2}x\Big)$$

We assume that the events that we measure are located in the moving frame where τ and ξ are coordinates. We see that

$$\frac{\mathrm{d}\tau}{\mathrm{d}t} = \gamma\Big(1 - \frac{v}{c^2}\frac{\mathrm{d}x}{\mathrm{d}t}\Big) = \gamma\Big(1 - \frac{v^2}{c^2}\Big) = \frac{\gamma}{\gamma^2} = \frac{1}{\gamma}$$

$$\therefore$$

$$\frac{\mathrm{d}\tau}{\mathrm{d}t} = \sqrt{1 - \frac{v^2}{c^2}} \to \mathrm{d}\tau^2 = \mathrm{d}t^2 - \frac{v^2\mathrm{d}t^2}{c^2} \to$$

$$c^2\mathrm{d}\tau^2 = (c\mathrm{d}t)^2 - \mathrm{d}x^2 = g_{ij}\mathrm{d}x^i\mathrm{d}x^j$$

where $i, j = 0, 1$, $\mathrm{d}x^0 = c\mathrm{d}t$, and $\mathrm{d}x^1 = \mathrm{d}x$. The 2×2 matrix in g_{ij} is

$$g_{ij} = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$$

As we see, $\mathrm{d}s$ is not the module of the vector $(c\mathrm{d}t, \mathrm{d}x)$ under the normal scalar product, which will correspond to a g_{ij} equal to the unit matrix

$$\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$

We may see it, however, as the module of the vector $(c\mathrm{d}t, i\mathrm{d}x)$ with an imaginary x component.

Relativity According to Minkowsky

The principals of Minkowsky Space are:

1. Any object or ”substance” in space and time is a point in the 4D world. The evolution of that point describes a world-line in the space. Objects that are stationary or at rest in the (x, y, z, t) coordinate system have a world line parallel to Ot. Objects with constant velocity v have a world-line forming a constant angle with Ot, such that $\frac{\mathrm{d}x}{\mathrm{d}t} = v$ indicates the slope of a straight line. A world-line parallel to Ox or perpendicular to Ot indicate an instant view of the world.

2. A very important scalar is the magnitude of a vector, which represents the distance between events. In orthogonal coordinates, this is expressed as

   $$c^2(t - t_0)^2 - (x - x_0)^2 - (y - y_0)^2 - (z - z_0)^2 = k^2 \quad (0.7.26)$$

   The value of k^2 is independent of the coordinate system that may be changed by any linear transformation, or equivalently by any rotation of the 4D system. With $k^2 = 0$, the scalar represents the transmission of light with velocity c from P_0 to P or vice versa. In general, the scalar product of a vector in the four dimensions is

   $$\vec{W} \cdot \vec{W} = x^i y^i g_{ij} \quad (0.7.27)$$

   Comparing (0.7.26) and (0.7.27), the matrix g_{ij} for orthogonal coordinates is

   $$g_{ij} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix}$$

   In other systems of reference (for instance, in spherical coordinates), g_{ij} depends on the coordinates.

A point in the space-time $P(x, y, z, t)$ may describe a line that we call a "world-line." As it moves from P to $P + \delta P$ (a point near P) with velocity $v < c$, the value of the differential distance in the space-time is (assuming two-dimensions only)

$$\mathrm{d}s^2 = c^2 \mathrm{d}t^2 - \mathrm{d}x^2$$

If $\mathrm{d}s = 0$, the substance described by $P(x, t)$ is a photon moving with velocity c; otherwise, the substance moves with velocity $v = \frac{\mathrm{d}x}{\mathrm{d}t} < c$. The increment $\mathrm{d}s$ is greater than zero and invariant with respect to a change in the coordinates, say $t \to t'$, or $x \to x'$.

"Proper time" τ is given by the following equation:

$$c^2 \mathrm{d}\tau^2 = \mathrm{d}s^2 = c^2 \mathrm{d}t^2 - \mathrm{d}x^2$$

For stationary points, $\mathrm{d}\tau = 0$ and

$$\mathrm{d}\tau = \mathrm{d}t \quad , \quad \tau = t$$

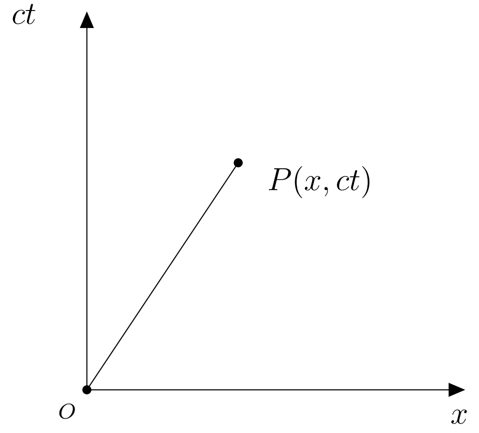
That is, the proper time is the same as t. If the substance is at rest in the system (t', x') that moves with velocity v, the proper time is t'.

The invariant $\mathrm{d}s$ may be seen as the distance in space-time (clock and ruler) of two close events. and if the two events are simultaneous in the reference system (x, t), $\mathrm{d}s^2 = 0$ and it will be so in any other system related to (x, t) by linear transformation.

When the system of coordinates is not orthogonal, then

$$\mathrm{d}s^2 = g_{ij} \mathrm{d}x^i \mathrm{d}x^j$$

where g_{ij} is a function of the coordinates, $i, j = 1, 4$.



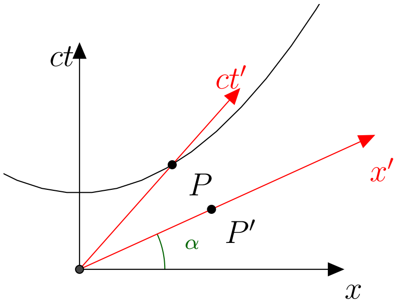
The point P is a point in the space-time coordinates. If a photon leaves O and reaches P in the direction OP, it will be at P at the (x, t) space-time point. The coordinates of P are such that

$$(ct)^2 - x^2 = 0$$

A mobile with velocity $v < c$ follows the same world trajectory $\overrightarrow{OP}$. The coordinates at P will fulfill an equation such that $\overrightarrow{OP} \cdot \overrightarrow{OP} = a^2$, or

$$(ct)^2 - x^2 = a^2$$

where a is a constant greater than zero. Taking $a = 1$ and allowing P to move around O, $\overrightarrow{OP}$ will describe all the objects that leave O and reach P with velocity $\dfrac{\mathrm{d}x}{\mathrm{d}t} = v$. The point P describes a hyperbola $c^2t^2 = x^2 + 1$. If we choose $\overrightarrow{OP}$ as the direction of the new time coordinates, then those objects will be at rest in the new system with time coordinate t'.



Once we choose the axis Ot' we choose an orthogonal axis, Ox', such that the inner product $\overrightarrow{OP} \cdot \overrightarrow{OP'} = 0$. The dot product in this case will have the following g matrix:

$$g = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$$

so

$$c^2tt' - xx' = 0$$

Since $\overrightarrow{OP} = (ct, x)$, then $\overrightarrow{OP'} = (x, \sqrt{1+x^2})$.

It is simple to see that the tangent to the hyperbola at P is also perpendicular to $\overrightarrow{OP}$. In fact, the tangent at P has a slope α such that

$$\tan\alpha = \frac{\mathrm{d}(ct)}{\mathrm{d}x} = \frac{x}{\sqrt{1+x^2}}$$

because $ct = \sqrt{1+x^2}$. Then

$$\cos\alpha = \frac{1}{\sqrt{1+\tan^2\alpha}} = \frac{\sqrt{1+x^2}}{\sqrt{1+2x^2}}$$

and

$$\sin\alpha = \frac{\tan\alpha}{\sqrt{1+\tan^2\alpha}} = \frac{x}{\sqrt{1+2x^2}}$$

may be taken as the components of the direction $\overrightarrow{OP'}$ along $\overrightarrow{Ox'}$.

$$\overrightarrow{OP'} = (\cos\alpha, \sin\alpha) = \left(\frac{\sqrt{1+x^2}}{\sqrt{1+2x^2}}, \frac{x}{\sqrt{1+2x^2}}\right)$$

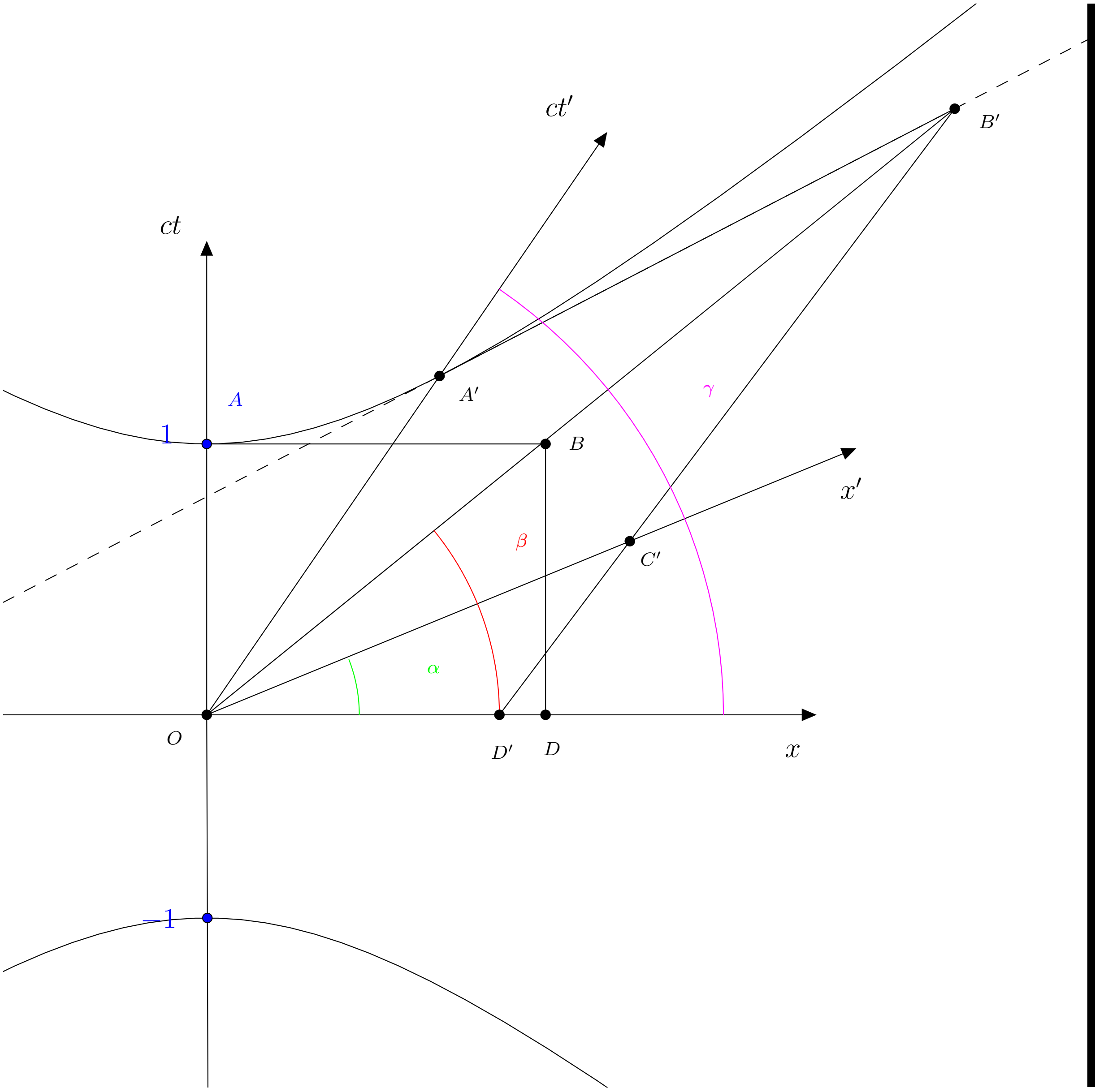
And thus, $\overrightarrow{OP'}$, which is parallel to the tangent, is perpendicular to $\overrightarrow{OP}$, which was taken as the time t' axis. Explicitly, $\overrightarrow{OP} = (ct, x)$ and

$$x\cos\alpha = \sqrt{1+x^2}\sin\alpha$$

or

$$\overrightarrow{OP'} \cdot \overrightarrow{OP} = \frac{x\sqrt{1+x^2}}{\sqrt{1+2x^2}} - \frac{x\sqrt{1+x^2}}{\sqrt{1+2x^2}} = 0$$

Lorentz Length Contraction From the Geometry of World-Lines



Let us return to the hyperbola in orthogonal coordinates lying in the (t, x) plane and given by the equation $c^2t^2 - x^2 = 1$ or

$$t = \pm\frac{1}{c}\sqrt{1+x^2}$$

The curve consists of two different hyperbolas, one for $t > 0$ and the other for $t < 0$. They are both symmetric with respect to the origin O. The asymptote passes through $B(1, 1)$ and OD is taken as 1. In fact, the asymptotes have an equation given by

$$c\frac{\mathrm{d}t}{\mathrm{d}x} = \pm\frac{x}{\sqrt{1+x^2}}_{x\to\infty} = \pm 1$$

Thus, the equation of the asymptote is

$$t = \pm\frac{1}{c}x$$

and the so the point B is on the asymptote.

The tangents of the angles α, β, and γ are such that

$$\tan\alpha = \frac{v}{c^2} \quad , \quad \tan\beta = 1 \quad , \quad \tan\gamma = \frac{c}{v}$$

$\tan\gamma = \frac{v}{c}$ is the slope of the straight line Ot'. By choosing the new time coordinate of Ot' passing through A', the point A' will be at rest in this system of (x', t'). Next, we choose Ox' parallel to the tangent at A' that has the slope

$$\left(\frac{c\mathrm{d}t}{\mathrm{d}x}\right)_{A'} = \left(\frac{x}{\sqrt{1+x^2}}\right)_{A'}$$

As A' goes to A on the world curve, $\left(\frac{\mathrm{d}t}{\mathrm{d}x}\right)_{A'}$ is a horizontal line parallel to Ox. Also, as $c \to \infty$, $\left(\frac{\mathrm{d}t}{\mathrm{d}x}\right)_{A'}$ is zero, so in both curves Ox' becomes identical to Ox. As $x_{A'} \to \infty$, $\left(\frac{\mathrm{d}t}{\mathrm{d}x}\right)_{\infty} = \frac{1}{c}$, or $c\frac{\mathrm{d}t}{\mathrm{d}x} = 1$, which is the slope of the asymptote (the angle β).

The rhomboid $A'B'D'O$ becomes $OABD$ as A' becomes A, B' becomes B, and D' becomes D. OD' has the same length as OC' as seen by the observer at the (x, t) system. The ratio $\frac{OD'}{OD}$ indicates the difference of length of a ruler OD as (x', t') moves with velocity v. So the same ruler is seen by the observer at (x, t) once at rest

(OD) and then moving with velocity v as OD'. If the ruler is at rest on Ox', its world-line is parallel to Ot' and its slope is $\dfrac{1}{v}$. So, $\tan\alpha = \dfrac{1}{v}$. Since $\dfrac{\mathrm{d}x}{\mathrm{d}t} = v$, v is the tangent of the angle between Ot and Ot'.

To find the ratio between OD' and OD, we use the following analytic procedure, our first goal being to find the coordinates of B'.

$$\tan\alpha = \left(c\frac{\mathrm{d}t}{\mathrm{d}x}\right)_{A'} = \left(\frac{x}{\sqrt{1+x^2}}\right)_{A'}$$

Since A' is on Ot', $\dfrac{t_{A'}}{x_{A'}} = \dfrac{1}{v}$, and since it is also on the hyperbola $c^2t^2_{A'} = x^2_{A'} + 1$.

$$x_{A'} = \frac{1}{\sqrt{\dfrac{c^2}{v^2} - 1}} \quad \text{and} \quad t_{A'} = \frac{1}{c\sqrt{1 - \dfrac{v^2}{c^2}}}$$

Substituting $x_{A'}$ into the equation for $\tan\alpha$ we find that

$$\tan\alpha = \frac{v}{c}$$

The point B' is the intersection of the asymptote $t_{B'} = \dfrac{1}{c}x_{B'}$ and the line $\overrightarrow{D'B'}$, whose slope is $\tan\gamma = \dfrac{c}{v}$. So the value of the coordinates at B' are

$$t_{B'} = \frac{1}{v}(x_{B'} - OD')$$

and so we get

$$x_{B'}\left(\frac{1}{v} - \frac{1}{c}\right) = \frac{1}{v}OD'$$

or

$$OD' = \left(1 - \frac{v}{c}\right)x_{B'}$$

We see that as $v \to c$, $\overrightarrow{OD'} \to 0$.

Also, since B' is on the tangent at A', whose slope is $\tan\alpha = \dfrac{v}{c}$:

$$c(t_{B'} - t_{A'}) = v(x_{B'} - x_{A'})$$

Substituting $t_{B'}$, $t_{A'}$, and $x_{A'}$ into the equation with the values previously found,

$$x_{B'} = \frac{\sqrt{1 - \frac{v^2}{c^2}}}{1 - \frac{v}{c}}$$

Thus,

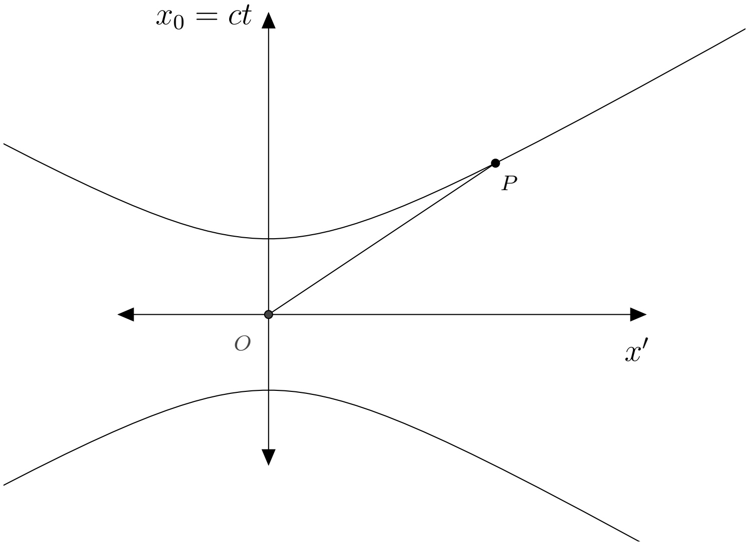
$$OD' = \sqrt{1 - \frac{v^2}{c^2}}$$

and since OD was taken as 1,

$$\frac{OD'}{OD} = \sqrt{1 - \frac{v^2}{c^2}}$$

This means that as the ruler moves with velocity v on the Ox axis, it is seen as shorter by a factor of $\sqrt{1 - \frac{v^2}{c^2}}$. This was first discovered by Lorentz while he was working on Electron Theory, although he assumed a unique and absolute time for every event in the world and for every observer. Einstein was the first to take time as a new variable: a different time-clock for every observer.

Proper Time as the Parameter to Create Velocity and Acceleration in the Space-Time System of Coordinates



Since $x^{0^2} - x^{1^2} = 1$,

$$2x^0 \frac{\mathrm{d}x^0}{\mathrm{d}s} - 2x^1 \frac{\mathrm{d}x^1}{\mathrm{d}s} = 0 \ .$$

Calling $\frac{\mathrm{d}x^0}{\mathrm{d}s} = x^0$ and $\frac{\mathrm{d}x^1}{\mathrm{d}s} = x^1$,

$$x^0 \dot{x}^0 - x^1 \dot{x}^1 \ .$$

Thus, (x^0, x^1) is orthogonal to $(\dot{x}^0, \dot{x}^1)$ according to the new inner product discussed above. Taking the derivative of the previous equation,

$$\dot{x}^{0^2} + \ddot{x}^0 x^0 - \dot{x}^{1^2} - \ddot{x}^1 x^1 = 0$$

or

$$\ddot{x}^0 x^0 - \ddot{x}^1 x^1 + \dot{x}^{0^2} - \dot{x}^{1^2} = 0$$

$$\ddot{x}^0 x^0 - \ddot{x}^1 x^1 = c^2 - v^2 \ .$$

Let us call $\dot{x}^{0^2} - \dot{x}^{1^2} = v^2$ the magnitude of the time-space velocity squared. In this simple case, it is $(c^2 - v^2)$. The vector acceleration in this space is $(\ddot{x}^0, \ddot{x}^1) - A$ such that

$$(\ddot{x}^0, \ddot{x}^1)(x^0, x^1) = v^2 \ .$$

Taking the magnitude of these vectors, the acceleration is

$$A = |\ddot{x}^0, \ddot{x}^1| = \frac{v^2}{|x^0, x^1|} = \frac{v^2}{\rho}$$

where ρ is the magnitude of the vector $\overrightarrow{OP}$ or the radius of curvature.

For any world-line passing through P, this is a hyperbola with the same tangent at P (or with 3 points in common). The curvature of that hyperbola is then equal to that of the world-line at P.

If the acceleration $(\ddot{x}^0, \ddot{x}^1)$ is zero, then the radius of curvature is ∞ and the world-curvature is a straight line.

Relativistic Dynamics: Application of Tensor Theory to Special Relativity

Any meaningful quantity in the theory of relativity must vary covariantly under coordinate transformation in a geometry generated by the inner product consistent with the Lorentz transformation. The fundamental tensor of such a geometry, $g_{\mu\nu}$, in the four-dimensional space x^μ (where $\mu = 0, 1, 2, 3$ and $x^0 = ct, x^1 = x, x^2 = y, x^3 = z$ in rectilinear coordinates) is such that

$$\mathrm{d}s^2 = g_{\mu\nu}\mathrm{d}x^\mu \mathrm{d}x^\nu$$

is invariant.

Again for rectilinear coordinates,

$$\mathrm{d}s^2 = c^2 t^2 - \mathrm{d}x^2 - \mathrm{d}y^2 - \mathrm{d}z^2$$

or

$$g_{\mu\nu} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix}$$

To make our discussion simple, let us work in rectilinear coordinates and two dimensions only, so $x^0 = ct$ and $x^1 = x$. Thus, if the velocity of the particle $\frac{\mathrm{d}x}{\mathrm{d}t} = v$ is constant, then

$$\mathrm{d}s^2 = (c^2 - v^2)\mathrm{d}t^2$$

or

$$\mathrm{d}s = \sqrt{c^2 - v^2}\mathrm{d}t$$

After integrating and squaring again with $s(t = 0) = 0$, we have

$$s^2 = c^2 t^2 - x^2$$

Differentiating with respect to t, we find that

$$\vec{v} \cdot \vec{x} = 0$$

where $\vec{x}$ is the time-space position vector of the point M in study, and $\vec{v}$ is the velocity in the O frame where the observer measured t. Thus, in general we may say that the four vector position and its four vector velocity in the S frame are orthogonal, by the inner product determined by $g_{\mu\nu}$.

By taking the second derivative with respect to t of $s^2 = {x^0}^2 - {x^1}^2$ we find that

$$\vec{a} \cdot \vec{v} = -\frac{c^2}{\gamma^2}$$

where $\vec{a}$ and $\vec{v}$ are the acceleration and velocity of the point M in the O frame, where the observer measures the time t. Thus,

$$\vec{a} = (\ddot{x}^0, \ddot{x}^1) \quad \text{and} \quad \vec{v} = (\dot{x}^0, \dot{x}^1)$$

where the dot means the derivative with respect to t.

We observe that the vectors $\vec{a}$ and $\vec{v}$ depend on the frame of reference. However, the following definitions of velocity and acceleration are tensors that depend on the particle itself only.

Given the four-vector displacement $x^0, x^1, x^2, x^3 : x^\mu$ as a function of the proper time τ (which is measured by the clock at the particle in question), the following derivatives form a vector that we call the velocity of the particle:

$$U^\mu = \frac{\mathrm{d}x^\mu}{\mathrm{d}\tau}$$

In the space-time, $x^\mu(\tau)$ describes a curve with a differential length given by $\mathrm{d}s = c\mathrm{d}\tau$ where c is the velocity of light. If we wish to relate U^μ to the three component velocity $\vec{u}$ determined by the variation of x^1, x^2, and x^3 with respect to t, we must remember that there is a time dilation such that

$$\mathrm{d}t = \gamma(u)\mathrm{d}\tau$$

Thus,

$$\frac{\mathrm{d}x^\mu}{\mathrm{d}\tau} = \frac{\mathrm{d}x^\mu}{\mathrm{d}t}\frac{\mathrm{d}t}{\mathrm{d}\tau} = u^i\gamma(u)$$

where $i = 1, 2, 3$ and $\gamma(u) = \dfrac{1}{\sqrt{1 - \dfrac{u^2}{c^2}}}$.

We have shown that $\dfrac{\mathrm{d}x^\mu}{\mathrm{d}t} = (c, \vec{u})$, so

$$U^\mu = \gamma(u)(c, \vec{u})$$

and the covariant component of U is

$$U_\mu = \gamma(u)(c, -\vec{u})$$

The magnitude of U^2 is

$$U^2 = U^\mu U_\mu = \gamma^2(u)(c^2 - u^2) = c^2$$

Thus, U^2 is invariant. This was not the case of v^2 previously discussed.

The corresponding acceleration is defined as

$$A^\mu = \frac{\mathrm{d}U^\mu}{\mathrm{d}\tau}$$

Since u also depends on τ,

$$A^\mu = \frac{\mathrm{d}}{\mathrm{d}t}[\gamma(u)(c, \vec{u})] = \gamma(u)\Big[c\frac{\mathrm{d}\gamma}{\mathrm{d}t}, \frac{\mathrm{d}\gamma}{\mathrm{d}t}\vec{u} + \gamma\vec{a}\Big]$$

where

$$\vec{a} = \frac{\mathrm{d}^2}{\mathrm{d}t^2}(x^1, x^2, x^3) = \frac{\mathrm{d}^2}{\mathrm{d}t^2}(x, y, z)$$

which is the three-vector acceleration measured by the observer whose clock reads time t.

The covariant A_μ is equal to

$$A_\mu = g_{\mu\nu}A^\nu = \gamma(u)\Big[c\frac{\mathrm{d}\gamma}{\mathrm{d}t}, -\Big(\frac{\mathrm{d}\gamma}{\mathrm{d}t}\vec{u} + \gamma\vec{a}\Big)\Big]$$

The expression of A^μ may be simplified in the following cases:

1. If $u = 0$, i.e., we are measuring A in the instantaneous rest frame, then

   $$\gamma(0) = 1 \quad \text{and} \quad \frac{\mathrm{d}\gamma}{\mathrm{d}t} = 0$$

   so

   $$A^\mu = (0, a) \quad \text{and} \quad A_\mu = A^\nu g_{\mu\nu} = (0, -a)$$

   The proper acceleration vector $\alpha = a$, and the scalar A^2 is

   $$A^2 = A_\mu A^\mu = -a^2$$

2. If the magnitude of $\vec{u}$ does not vary with time, such as in uniform circular motion where $a = \dfrac{|u|^2}{r}$, then $\dfrac{\mathrm{d}\gamma}{\mathrm{d}t} = 0$ and

   $$A^\mu = \gamma(u)(0, \gamma a)$$

   and the scalar $A^2 = A^\mu A_\mu = \gamma^2(u)(0, -\gamma^2 a^2)$, so

   $$\alpha = \gamma^2 a = \frac{\gamma^2 |u|^2}{r}$$

Relativistic Forces

The four-vector force is defined as

$$F = \frac{\mathrm{d}}{\mathrm{d}\tau}(m_0 U)$$

where m_0 is the mass of the particle in its own frame. Let us relate F to the classic Newtonian force f:

$$\begin{aligned} F = \gamma(u)\frac{\mathrm{d}}{\mathrm{d}t}(m_0\gamma c, m_0\gamma\vec{u}) = \gamma(u)\frac{\mathrm{d}}{\mathrm{d}t}(mc, mu) &= \gamma(u)\Big(\frac{\mathrm{d}(mc^2)}{c\mathrm{d}t}, \frac{\mathrm{d}}{\mathrm{d}t}\vec{p}\Big) \\ = \gamma(u)\Big(\frac{1}{c}\frac{\mathrm{d}E}{\mathrm{d}t}, \frac{\mathrm{d}\vec{p}}{\mathrm{d}t}\Big) &= \gamma(u)\Big(\frac{1}{c}\frac{\mathrm{d}E}{\mathrm{d}t}, f\Big)\ . \end{aligned}$$

Using tensor notation, we may find the invariant $I = F^\mu U^\nu g_{\mu\nu}$, where

$$F^\mu = (F^0, F^1, F^2, F^3) = \gamma(u)\Big(\frac{\mathrm{d}E}{\mathrm{d}x^0}, f^1, f^2, f^3\Big)$$

and

$$U^\nu = (U^0, U^1, U^2, U^3) = \gamma(u)(c, \vec{u})\ .$$

Then,

$$\begin{aligned} I = F \cdot U = F^\mu U^\nu g_{\mu\nu} &= \gamma(u)\Big(\frac{1}{c}\frac{\mathrm{d}E}{\mathrm{d}t}, \vec{f}\Big) \cdot \gamma(u)(c, \vec{u}) g_{\mu\nu} \\ &= \gamma^2(u)\Big(\frac{\mathrm{d}E}{\mathrm{d}t}, -\vec{f} \cdot \vec{u}\Big) \end{aligned}$$

This invariant I, at the rest frame where $m = m_0$, $\gamma(u) = 1$, and $U = (c, 0)$, is

$$I = c^2\frac{\mathrm{d}(m_0)}{\mathrm{d}t} = c^2\Big(\frac{\mathrm{d}m_0}{\mathrm{d}\tau}\Big)$$

since $\mathrm{d}t = \mathrm{d}\tau$ for the rest frame. Equating both Is,

$$I = c^2\frac{\mathrm{d}m_0}{\mathrm{d}\tau} = \gamma^2(u)\frac{\mathrm{d}E}{\mathrm{d}t} - \gamma^2(\vec{f} \cdot \vec{u})\ .$$

A pure force means that there was no change in the rest mass m_0. Thus, for a pure force ($\frac{\mathrm{d}m_0}{\mathrm{d}t} = 0$), then $I = 0$ and

$$\frac{\mathrm{d}E}{\mathrm{d}t} = (\vec{f} \cdot \vec{u})$$

and so $I = F \cdot U = 0$. The force F, when it is a pure force, can be written as

$$\vec{F} = \gamma(u)\Big(\frac{\vec{f} \cdot \vec{u}}{c}, \vec{f}\Big) \ .$$

The following examples may prove helpful in understanding the difference between the three-vector and the four-vector velocity and acceleration:

Let us look at a high-velocity particle in an accelerator where the path is a circular one with radius of curvature R at the moment when we study the state of the particle. If the instantaneous three-vector velocity is along the path, then there will be a radial acceleration equal to $\frac{u^2}{R}$ that is normal to the velocity. All of this is from the reference frame of the observer at the accelerator. We then ask what the acceleration is for the observer at the particle, that is, the proper acceleration A^μ in the four time-space coordinates, and $\vec{\alpha}$, the three-vector proper acceleration.

$$A^2 = A^\mu A_\mu = -\alpha^2 \ .$$

Now

$$A^\mu = \gamma(u)\Big[c\Big(\frac{\mathrm{d}\gamma}{\mathrm{d}t}\Big), \gamma\vec{a} + \vec{u}\Big(\frac{\mathrm{d}\gamma}{\mathrm{d}t}\Big)\Big] \ .$$

If we assume that $|\vec{u}|$ is constant in time, then $\frac{\mathrm{d}\gamma}{\mathrm{d}t}$ (which implies $\frac{\mathrm{d}u}{\mathrm{d}t}$) will be zero, and

$$A^\mu = \gamma(u)[0, \gamma\vec{a}]$$

and

$$A^2 = A^\mu A_\mu = -\gamma^4(u)a^2 \ .$$

So

$$\alpha = \gamma^2(u)a \ .$$

For example, if $u = 0.888c$, then

$$\alpha = \frac{1}{1-(0.888)^2} \times (0.888)^2\frac{c^2}{R^2} = 4.2\frac{c^2}{R^2} \ .$$

It is common to express the Lorentz transformation using the hyperbolic functions. For instance, if we define the angle Φ (sometimes called the rapidity) as

$$\tanh\Phi = \frac{u}{c}$$

where u is the magnitude of $\vec{u}$, then,

$$\frac{\mathrm{d}}{\mathrm{d}t}(\tanh\Phi) = \frac{1}{c}\frac{\mathrm{d}u}{\mathrm{d}t} = \frac{a}{c} \ .$$

But the first member is

$$\left(\frac{1}{\cosh^2}\right)\frac{\mathrm{d}\Phi}{\mathrm{d}t} = (1-\tanh^2\Phi)\frac{\mathrm{d}\Phi}{\mathrm{d}t} = \left(1-\frac{u^2}{c^2}\right)\frac{\mathrm{d}\Phi}{\mathrm{d}t} = \frac{1}{\gamma^2}\frac{\mathrm{d}\Phi}{\mathrm{d}t} \ .$$

If the frame (τ,ξ,η,ζ) sees a particle in rectilinear motion, then its acceleration a as seen by S, $(\frac{\mathrm{d}u}{\mathrm{d}t})$, is related to $\alpha = \frac{\mathrm{d}u'}{\mathrm{d}\tau}$ by $\alpha = \gamma^3 a$, where the primes refer to the coordinates in the moving frame. Thus,

$$\frac{1}{\gamma^2}\frac{\mathrm{d}\Phi}{\mathrm{d}t} = \frac{\alpha}{c}\frac{1}{\gamma^3}$$

and

$$\gamma\frac{\mathrm{d}\Phi}{\mathrm{d}t} = \frac{\alpha}{c}$$

but $(\mathrm{d}t)\frac{1}{\gamma} = \mathrm{d}\tau$, so

$$\frac{\mathrm{d}\Phi}{\mathrm{d}\tau} = \frac{\alpha}{c} \ .$$

If α is constant with adequate boundary conditions, then

$$\Phi = \frac{\tau\alpha}{c} \ .$$

In this particular case of α being constant, we find interesting relationships:

1. Since $\tanh\Phi = \frac{u}{c}$, then

$$\frac{u}{c} = \tanh\left(\frac{\tau\alpha}{c}\right) \ .$$

2.
$$\gamma(u) = \frac{1}{\sqrt{1-\tanh^2\Phi}} = \cosh\Phi = \cosh\left(\frac{\tau\alpha}{c}\right)$$

3. In a previous discussion, we found that

$$\vec{\alpha} = \frac{\mathrm{d}\vec{u}'}{\mathrm{d}t'} = \frac{\mathrm{d}}{\mathrm{d}t}(\gamma(u)\vec{u})$$

which is also true for the magnitude of $\vec{u}$ and $\vec{u}'$, so we may write

$$\alpha = \frac{\mathrm{d}}{\mathrm{d}t}(\gamma(u)u)$$

and since α is a constant,

$$t\alpha = \gamma(u)u \tag{0.7.28}$$

with an adequate boundary condition of $t = 0$. Thus, using the previous derivation for $\gamma(u)$,

$$\frac{t\alpha}{c} = \gamma(u)\frac{u}{c} = \cosh(\Phi)\tanh\Phi = \sinh\Phi = \sinh\left(\frac{\tau\alpha}{c}\right) .$$

This gives an equation that relates the times t and τ and the value of α:

$$\frac{t\alpha}{c} = \sinh\left(\frac{\tau\alpha}{c}\right) . \tag{0.7.29}$$

4. Since $u = \dfrac{\mathrm{d}x}{\mathrm{d}t}$, then

$$\frac{\mathrm{d}x}{\mathrm{d}t} = u = c\tanh\Phi .$$

Integrating and assuming some x_0 when $t = 0$ (also assuming that $\tau = 0$ for $t = 0$),

$$\int_{x_0}^{x} \mathrm{d}x = (x - x_0) = c\int_{t=0}^{t} \tanh(\Phi)\ \mathrm{d}t = c\int_0^{\tau} \gamma\tanh\Phi\ \mathrm{d}\tau .$$

From the expression of γ previously derived,

$$x - x_0 = c\int_0^{\tau} \sinh\Phi\ \mathrm{d}\tau = \int_{\Phi_0}^{\Phi} \frac{c^2}{\alpha}\sinh\Phi\ \mathrm{d}\Phi$$

or

$$x - x_0 = \frac{c^2}{\alpha}\Big[\cosh\Phi\Big]_{\Phi_0}^{\Phi} \ .$$

Since $\tau = 0$ for $t = 0$, $\Phi_0 = 0$. Then,

$$x = x_0 + \frac{c^2}{\alpha}\cosh\Phi \ .$$

Let us define another property of Φ, the concept of "rapidity" of a particle moving with respect to O_1 with velocity

$$u = \frac{u' + v}{1 - \dfrac{u'v}{c^2}}$$

where u' is the relative velocity of O_2 that moves with respect to O_1 with velocity v. The "rapidity" of Φ is such that its hyperbolic tangent is equal to $\frac{u}{c}$:

$$\tanh\Phi = \frac{u}{c} \ .$$

If the particle is at the observer O_2, $u' = 0$ and $u = v$, thus

$$\tanh\Phi = \frac{v}{c} \ .$$

Assume now that there are three observers: O_1, O_2, O_3, and three particles located at the observers' positions. If v_1 is the relative velocity of O_2 with respect to O_1, then the "rapidity" of the particle at O_2 with respect to O_1 is Φ such that

$$e^{\Phi} = \sqrt{\frac{c + v_1}{c - v_1}}$$

This equation is easily derivable using the expression for $\sinh\Phi$ and $\cosh\Phi$, and by recalling that

$$\sinh\Phi + \cosh\Phi = e^{\Phi} \ .$$

If Ψ is the rapidity of the particle at O_3 with respect to the observer at O_2, then

$$e^{\Psi} = \sqrt{\frac{c+v_2}{c-v_2}}$$

where v_2 is the velocity of the particle at O_3 with respect to that at O_2.

The velocity of the particle at O_3 with respect to that of the particle at O_1 is

$$u = \frac{u'+v_1}{1+\dfrac{u'v_1}{c^2}}$$

where u' is the relative velocity of the particle at O_3 with respect to O_2, and is equal to v_2. Thus,

$$u = \frac{v_1+v_2}{1+\dfrac{v_1v_2}{c^2}}$$

and the rapidity θ of the particle at O_3 with respect to O_1 is

$$e^{\theta} = \sqrt{\frac{c+u}{c-u}}$$

Substituting u into the expression above,

$$e^{\theta} = \left(\left(c+\frac{v_1+v_2}{1+\dfrac{v_1v_2}{c^2}}\right)\div\left(c-\frac{v_1+v_2}{1+\dfrac{v_1v_2}{c^2}}\right)\right)^{\frac{1}{2}} = \sqrt{\frac{c^2+c(v_1+v_1)+v_1v_2}{c^2-c(v_1+v_2)+v_1v_2}}$$ [redacted]

As we can easily see, this is equal to the product

$$e^{\Phi}\cdot e^{\Psi} = \sqrt{\frac{c+v_1}{c-v_1}}\sqrt{\frac{c+v_2}{c-v_2}}$$

So the rapidity θ of the particle at O_3 with respect to the observer at O_1 is the sum of the rapidity of the particle at O_3 with respect to O_2 and the particle at O_2 with respect to O_1. Thus,

$$\theta = \Phi + \Psi\ .$$

If instead of three particles we have n particles with rapidities with respect to each other, $\Phi_1, \Phi_2, \Phi_3, \ldots, \Phi_n$, the rapidity of the nth particle with respect to O_1 is θ such that

$$\theta = \Phi_1 + \Phi_2 + \cdots + \Phi_n \ .$$

Calling v the velocity of the nth particle with respect to O_1, we may write

$$e^\theta = \sqrt{\frac{c+v}{c-v}} \ .$$

Assuming $\Phi_1 = \Phi_2 = \cdots = \Phi_n = \Phi$,

$$e^\theta = e^{n\Phi}$$

so

$$e^{2\theta} = e^{2n\Phi} = \frac{c+v}{c-v} \ .$$

Calling $z^n = e^{2n\Phi}$,

$$\frac{v}{c} = \frac{z^n - 1}{z^n + 1} \ .$$

If we call v_0 the velocity corresponding to the rapidity Φ,

$$e^{2\Phi} = \frac{c+v_0}{c-v_0}$$

and

$$z^n = \left(\frac{c+v_0}{c-v_0}\right)^n \ .$$

Assuming that $v_0 = \frac{1}{2}c$, then

$$\frac{v}{c} = \frac{z^n - 1}{z^n + 1} = \frac{3^n - 1}{3^n + 1}$$

For $n = 5$, the particle in the proper system will acquire a speed

$$\frac{v}{c} = \frac{243 - 1}{243 + 1} = 0.992$$

Relativistic Lagrangian and the Lorentz Force

The momentum of a free particle is (we work in Cartesian coordinates):

$$p = m\dot{x} = \frac{m_0}{\sqrt{1-\beta^2}}\dot{x} \quad \text{where} \quad \dot{x} = \frac{\mathrm{d}x}{\mathrm{d}t} \tag{0.7.30}$$

The Lagrangian is related to p by

$$p = \frac{\partial \mathcal{L}_p}{\partial \dot{x}} \quad \text{where} \quad \mathcal{L}_p \text{ is the free particle Lagrangian}$$

By equating both expressions, we derive the following for the free particle Lagrangian:

$$\mathcal{L}_p = \int m_0\dot{x}\sqrt{1-\frac{\dot{x}^2}{c^2}}\,d\dot{x} = -m_0c^2\sqrt{1-\beta^2} \tag{0.7.31}$$

If we now assume an electromagnetic field interacting on the particle, then (the charge q is assumed > 0, the electron charge $e = -q$):

$$\mathcal{L} = -m_0c^2\sqrt{1-\beta^2} - q\left[\Phi - \frac{\vec{v}}{c}\cdot\vec{A}\right] \tag{0.7.32}$$

where Φ, the scalar potential, and $\vec{A}$, the vector potential, are both functions of x, y, z, t.

The Lagrange Equation is:

$$\frac{\mathrm{d}}{\mathrm{d}t}\frac{\partial \mathcal{L}}{\partial \dot{x}} - \frac{\partial \mathcal{L}}{\partial x} = 0$$

Let us separate the Lagrangian in the free particle and the interacting one:

$$\mathcal{L} = \mathcal{L}_p + \mathcal{L}_i$$

Note: In Classical Mechanics, the Lagrange is $\mathcal{L} = T - V$, where T is the kinetic energy and V is the potential energy. For the free particle Lagrangian, $\mathcal{L}_p$ will be equal to T which is > 0. In relativistic cases $\mathcal{L}_p = -m_0c^2\sqrt{1-\beta^2}$ is negative. To see that these facts are consistent, we expand $\mathcal{L}$ for low values of $v \ll c$. Let us assume β^2 to be the smallest increment in the expansion:

$$\mathcal{L}_p = -m_0c^2 - m_0c^2\Big(\frac{-\frac{1}{2}}{\sqrt{1-\beta^2}}\Big)_{\beta^2=0}\beta^2 + \ldots\beta^4 + \ldots\beta^6 + \ldots$$

$$\mathcal{L}_p \approx -m_0c^2 + \frac{m_0v^2}{2} + \ldots$$

Ignoring the rest energy, which plays no role in Classical Mechanics,

$$\mathcal{L}_p \approx T, \text{ the classical kinetic energy equal to } \frac{1}{2}m_0v^2$$

End of Note

As is well known the Coulomb force is

$$\vec{F}_c = q\vec{E}$$

and the Lorentz force due to the magnetic field B is

$$\vec{F}_L = \frac{q}{c}(\vec{v} \times \vec{B})$$

where $\vec{v}$ is the velocity of the particle and q the charge particle. Next, we will derive this last force from the previous relativistic Lagrangian.

The interacting Lagrangian, when an EM field is present, is given by

$$\mathcal{L}_i = -q(\Phi - \vec{A} \cdot \frac{\vec{v}}{c})$$

Next, we will see that this expression for the interacting Lagrangian gives rise to the Lorentz Force.

If Q_i is the external force, and since

$$\frac{\partial \mathcal{L}_p}{\partial x} = 0$$

then

$$\frac{\mathrm{d}}{\mathrm{d}t}\frac{\partial \mathcal{L}_p}{\partial \dot{x}} = \frac{\partial \mathcal{L}_i}{\partial x} - \frac{\mathrm{d}}{\mathrm{d}t}\frac{\partial \mathcal{L}_i}{\partial \dot{x}} = Q_i$$

Now

$$\mathcal{L}_p = -m_0 c^2 \sqrt{1-\beta^2}$$

and

$$\frac{\partial \mathcal{L}_p}{\partial \dot{x}} = m_0 c^2 \frac{\beta \dfrac{\partial \beta}{\partial \dot{x}}}{\sqrt{1-\beta^2}}$$

since

$$\beta = \frac{v}{c} = \frac{\sqrt{\dot{x}^2+\dot{y}^2+\dot{z}^2}}{c}$$

and

$$\frac{\partial \beta}{\partial \dot{x}} = \frac{\dot{x}}{\beta c^2} \ .$$

Thus

$$\frac{\partial \mathcal{L}_p}{\partial \dot{x}} = \frac{m_0 \dot{x}}{\sqrt{1-\beta^2}}$$

and

$$\frac{\partial \mathcal{L}}{\partial x} = \frac{\partial \mathcal{L}_i}{\partial x} = -q\Big[\frac{\partial \Phi}{\partial x} - \frac{\partial}{\partial x}\Big(\frac{\dot{x}}{c}A_x + \frac{\dot{y}}{c}A_y + \frac{\dot{z}}{c}A_z\Big)\Big] \ .$$

Also,

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}t}\frac{\partial \mathcal{L}_i}{\partial \dot{x}} &= \frac{\mathrm{d}}{\mathrm{d}t}\frac{\partial}{\partial \dot{x}}\Big(\frac{q}{c}\vec{v}\cdot\vec{A}\Big) = \frac{q}{c}\frac{\mathrm{d}}{\mathrm{d}t}\frac{\partial}{\partial \dot{x}}(\dot{x}A_x + \dot{y}A_y + \dot{z}A_z) = \frac{q}{c}\frac{\mathrm{d}}{\mathrm{d}t}A_x \\ &= \frac{q}{c}\Big(\frac{\partial A_x}{\partial t} + \frac{\partial A_x}{\partial x}\dot{x} + \frac{\partial A_x}{\partial y}\dot{y} + \frac{\partial A_x}{\partial z}\dot{z}\Big) \ .\end{aligned}$$

So, the Lagrange Equation becomes

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}t}\Big(\frac{m_0\dot{x}}{\sqrt{1-\beta^2}}\Big) + \frac{q}{c}\Big(\frac{\partial A_x}{\partial t} + \frac{\partial A_x}{\partial x}\dot{x} + \frac{\partial A_x}{\partial y}\dot{y} + \frac{\partial A_x}{\partial z}\dot{z}\Big) \\ = -q\frac{\partial \Phi}{\partial x} + \frac{q}{c}\Big(\frac{\partial A_x}{\partial x}\dot{x} + \frac{\partial A_x}{\partial x}\dot{y} + \frac{\partial A_x}{\partial x}\dot{z}\Big)\end{aligned}$$

or

$$\frac{\mathrm{d}}{\mathrm{d}t}\Big(\frac{m_0\dot{x}}{\sqrt{1-\beta^2}}\Big) = -\frac{q}{c}\frac{\partial A_x}{\partial t} - q\frac{\partial \Phi}{\partial x} + \frac{q}{c}\Big[\dot{y}\Big(\frac{\partial A_y}{\partial x} - \frac{\partial A_x}{\partial y}\Big) + \dot{z}\Big(\frac{\partial A_z}{\partial x} - \frac{\partial A_x}{\partial z}\Big)\Big] \ .$$ [redacted]

Thus,

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{m_0\dot{x}}{\sqrt{1-\beta^2}}\right) &= -\frac{q}{c}\frac{\partial A_x}{\partial t} - q\frac{\partial \Phi}{\partial x} + \frac{q}{c}(\dot{y}(\nabla\times A)_z + \dot{z}(\nabla\times A)_y) \\ &= -\frac{q}{c}\frac{\partial A_x}{\partial t} - q\frac{\partial \Phi}{\partial x} + \frac{q}{c}(v\times(\nabla\times A))_x\ .\end{aligned}$$

From the Maxwell Equations

$$E_x = -\frac{\partial \Phi}{\partial x} - \frac{\partial A_x}{c\partial t} \quad \text{and} \quad B = \nabla\times A$$

so finally

$$\frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{m_0\dot{x}}{\sqrt{1-\beta^2}}\right) = qE_x + \frac{q}{c}(\vec{v}\times\vec{B})_x \tag{0.7.33}$$

The last term represents the x-component of the Lorentz Force and the first component of the Coulomb Force. Thus,

$$Q_x = q\vec{E_x} + \frac{q}{c}(\vec{v}\times\vec{B})_x$$

and

$$\frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{m_0\dot{x}}{\sqrt{1-\beta^2}}\right) = Q_x\ .$$

Hamiltonian for a Charged Particle

The Hamiltonian is

$$\mathcal{H} = \mathcal{P}\dot{q}_i - \mathcal{L}$$

where (assuming Cartesian coordinates and dealing with x only)

$$\mathcal{P}_x = \frac{\partial \mathcal{L}}{\partial \dot{x}} = \frac{m_0 \dot{x}}{\sqrt{1-\beta^2}} + \frac{q}{c}A_x = m\dot{x} + \frac{q}{c}A_x = p_x + \frac{q}{c}A_x \ .$$

p_x is the free particle $(\sqrt{1-\beta^2})$ momentum. Thus, the first term of $\mathcal{H}$ is

$$\mathcal{P}\dot{x} = \frac{m_0 \dot{x}^2}{\sqrt{1-\beta^2}} + \frac{q}{c}A_x \dot{x}$$

and similarly for y and z. So,

$$\mathcal{P}_i \dot{q}_i = \frac{m_0 v^2}{\sqrt{1-\beta^2}} + \frac{q}{c}\vec{A}\cdot\vec{v} \ .$$

Since

$$\mathcal{L} = -m_0 c^2 \sqrt{1-\beta^2} - q(\Phi - \frac{\vec{v}}{c}\cdot\vec{A}) \ .$$

The reason for this is that $\vec{A}$ is the generator of $\vec{B}$ (the magnetic field), and it does not induce any work. Thus, $\mathcal{H}$ is the total energy, the sum of the kinetic mc^2 and the potential Φ.

The first term of $\mathcal{H}$ is just

$$\frac{m_0 v^2}{\sqrt{1-\beta^2}} + m_0 c^2 \sqrt{1-\beta^2}$$

or

$$\begin{aligned}\mathcal{H}_p^2 = \frac{m_0^2 c^4}{1-\dfrac{v^2}{c^2}} &= \frac{m_0^2 c^4 + m_0^2 c^2 v^2 - m_0^2 c^2 v^2}{1-\dfrac{v^2}{c^2}} \\ = \frac{m_0^2 c^2 (c^2 - v^2)}{1-\dfrac{v^2}{c^2}} &+ \frac{m_0^2 c^2 v^2}{1-\dfrac{v^2}{c^2}} = (m_0^2 c^4 + p^2 c^2)\end{aligned}$$

where $p = mv = \dfrac{m_0 v}{\sqrt{1-\beta^2}}$. So $\mathcal{H}$ may be expressed as a function of the kinetic momentum $p = mv$ such that

$$\mathcal{H} = E = \sqrt{(\vec{p}c)^2 + m_0^2 c^4} + q\Phi$$

or

$$\mathcal{H} = c\sqrt{m_0^2 c^2 + p_x^2 + p_y^2 + p_z^2} + q\Phi$$

where $m_0 c^2$ is the rest-mass energy, the addition to the classic Hamiltonian.

Now, using the definition of impulse of the coordinate x,

$$\mathcal{P}_x = \frac{\partial \mathcal{L}}{\partial \dot{x}}$$

we find that it is the sum of the kinetic momentum and potential,

$$\mathcal{P}_x = \frac{\partial \mathcal{L}}{\partial \dot{x}} = p_x + \frac{q}{c}\vec{A}$$

From this we find that

$$p_x = \mathcal{P}_x - \frac{q}{c}A_x$$

Substituting this into the expression for $\mathcal{H}$,

$$\mathcal{H} = c\sqrt{m_0^2 c^2 + (\mathcal{P} - \frac{q}{c}\vec{A})^2} + q\Phi = c\sqrt{m_0^2 c^2 + p^2} + q\Phi$$

For small velocities $v << c$, or $\dfrac{p^2}{m_0^2 c^2} << 1$, the Hamiltonian is

$$\mathcal{H} - m_0 c^2 = \frac{p^2}{2m_0} + q\Phi$$

The classical Hamiltonian is $\mathcal{H} - m_0 c^2$ (the constant $m_0 c^2$ does not show in its definition). Changing p for $\mathcal{P}$ as above, the classical Hamiltonian becomes

$$\mathcal{H} = \frac{1}{2m_0}\left(\mathcal{P} - \frac{q}{c}A\right)^2 + q\Phi$$

Note on Maupertuis and Fermat

Fermat (1601-1665) stated that an optical path is a minimum. If n is the index of refraction and ds the differential light path, then $\int_{M_0}^{M_1} n \, \mathrm{d}s$ is a minimum between the point M_0 and M_1.

Maupertuis (1698-1759) stated a similar principle in mechanics. It established that the following integral is a minimum:

$$\int_{M_0}^{M_1} p_i \dot{q}_i \, \mathrm{d}t$$

or in other words, the variation of that integral is zero, so

$$\delta \int_{M_0}^{M_1} p_i \dot{q}_i \, \mathrm{d}t = 0$$

Again, repeated indexes implies summation.

For a free particle, this means that

$$\int_{M_0}^{M_1} m\vec{v} \cdot \mathrm{d}\vec{s}$$

is minimized. (In relativity, $m \geq \dfrac{m_0}{\sqrt{1-\beta^2}}$). $\vec{v} \cdot \mathrm{d}\vec{s} = \dfrac{\mathrm{d}x}{\mathrm{d}t}\mathrm{d}x^i$ ($i = 1, 2, 3$) in Cartesian coordinates.

As we have seen, the generalized momentum is defined from the Lagrangian as

$$\mathcal{P}_i = \frac{\partial \mathcal{L}}{\partial \dot{q}_i}$$

The total Lagrange function for a charged particle in an electromagnetic field is given by the free-particle Lagrangian and the interactive one:

$$\mathcal{L} = \mathcal{L}_p + \mathcal{L}_Q$$

where

$$\mathcal{L}_p = -m_0 c^2 \sqrt{1-\beta^2}$$

and

$$\mathcal{L}_Q = -q\Phi + q\vec{A} \cdot \frac{\vec{v}}{c}$$

(q is the particle charge).

If we work with Cartesian coordinates, then the kinetic plus the momentum due to the potential $\vec{A}$ is

$$\mathcal{P}_x = \frac{q}{c}A_x + \frac{m_0\dot{x}}{\sqrt{1-\beta^2}} \quad , \quad \mathcal{P}_y = \dots$$

The second term of $\mathcal{P}$ may be written as a function of the Coulomb potential U if we assume that the particle's energy is conserved. So, the Hamiltonian may be expressed by

$$\mathcal{H} = mc^2 + qU = m_0c^2 + qU_0$$

Therefore,

$$(m - m_0)c^2 = q(U_0 - U)$$

but

$$m = \frac{m_0}{\sqrt{1 - \dfrac{v^2}{c^2}}}$$

Thus,

$$m_0c^2\left(\frac{1}{\sqrt{1-\dfrac{v^2}{c^2}}} - 1\right) = q(U_0 - U)$$

From this we obtain $\dfrac{m_0 v}{\sqrt{1-\dfrac{v^2}{c^2}}}$, which gives rise to the isotropic part of the path, while the unisotropy is due to the term $\dfrac{q}{c}\vec{A}$. After simple algebra, we obtain from the previous equation,

$$\mathcal{P}_{\text{isotropic}} = \frac{m_0\dfrac{v}{c}}{\sqrt{1-\dfrac{v^2}{c^2}}} = \sqrt{(U_0 - U)qm_0\left(\frac{q(U_0-U)}{m_0c^2} + 2\right)}$$

Canonical Equations of Hamilton–Poisson Brackets and Jacobi Algebra

From the Lagrangian we have defined the Hamiltonian $\mathcal{H}$ and the generalized momentum $\mathcal{P}$:

$$\mathcal{P}_i = \frac{\partial \mathcal{L}}{\partial \dot{q}_i} \quad \text{and} \quad \mathcal{H} = \mathcal{P}_i \dot{q}_i - \mathcal{L}$$

Assuming $\mathcal{P}_i$ and q_i are variables (called *canonical variables* by Jacobi),

$$\frac{\partial \mathcal{H}}{\partial \mathcal{P}_i} = \dot{q}_i \quad \text{and} \quad \frac{\partial \mathcal{H}}{\partial q_i} = -\frac{\partial \mathcal{L}}{\partial q_i} = -\frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\partial \mathcal{L}}{\partial \dot{q}_i}\right) = -\frac{\partial \mathcal{P}_i}{\partial t}$$

These are the *Canonical Equations* of Hamilton; that is,

$$\dot{q}_i = \frac{\partial \mathcal{H}}{\partial \mathcal{P}_i} \quad \text{and} \quad \dot{\mathcal{P}}_i = -\frac{\partial \mathcal{H}}{\partial q_i}$$

Given a function of q_i, $\mathcal{P}_i$, and t, $F(q_i, \mathcal{P}_i, t)$, its variation with respect to time is, using the Canonical Equations,

$$\begin{aligned} \frac{\mathrm{d}F}{\mathrm{d}t} &= \frac{\partial F}{\partial t} + \frac{\partial F}{\partial q_i}\dot{q}_i + \frac{\partial F}{\partial \mathcal{P}_i}\dot{\mathcal{P}}_i = \frac{\partial F}{\partial t} + \frac{\partial F}{\partial q_i}\frac{\partial \mathcal{H}}{\partial \mathcal{P}_i} - \frac{\partial F}{\partial \mathcal{P}_i}\frac{\partial \mathcal{H}}{\partial q_i} \\ &= \frac{\partial F}{\partial t} + \{F, \mathcal{H}\} = \frac{\partial F}{\partial t} + \text{Poisson Bracket of } F \text{ and } \mathcal{H} \end{aligned}$$

Jacobi defined the function $S(q_i, t)$ as

$$S(q,t) = -\int_{t_0}^{t} \mathcal{L}(q,t)\ \mathrm{d}t = \int_{t_0}^{t} \mathcal{H}\ \mathrm{d}t - \int_{t_0}^{t} \mathcal{P}_i\ \mathrm{d}q$$

From this we found that

$$\frac{\partial S}{\partial t} = \mathcal{H} \quad \text{and} \quad \frac{\partial S}{\partial q_i} = -\mathcal{P}_i$$

Since $\mathcal{H} = (\mathcal{P}_i \dot{q}_i - \mathcal{L})$, then substituting the two previous equations into this expression we obtain the Jacobi Equation:

$$\frac{\partial S}{\partial t} = -(\frac{\partial S}{\partial q_i}\dot{q}_i + \mathcal{L})$$

Observe that if we abstract the function S from these expressions, assuming $\mathcal{H}$ and $\mathcal{P}$ to be operators, we observe that $\mathcal{H}$ is associated with $\dfrac{\partial}{\partial t}$ and $\mathcal{P}$ with $-\dfrac{\partial}{\partial q}$. These correspondences are the basis of the Schrodinger Equations in Quantum Mechanics.

Poisson Brackets

Given a function $F(t, q, \mathcal{P}_i)$ in the phase space, i.e., the space of Canonical variables $q, \mathcal{P}_i$, its total derivative with respect to t is

$$\frac{\mathrm{d}F}{\mathrm{d}t} = \frac{\partial F}{\partial t} + \frac{\partial F}{\partial q}\frac{\mathrm{d}q}{\mathrm{d}t} + \frac{\partial F}{\partial \mathcal{P}_i}\frac{\mathrm{d}\mathcal{P}_i}{\mathrm{d}t} \ .$$

If the point $(q, \mathcal{P}_i)$ describes a dynamic imposed by a Hamiltonian $\mathcal{H}$, then as we have seen in lecture on Analytic Dynamics, the Hamiltonian Equations are:

$$\frac{\mathrm{d}\mathcal{P}_i}{\mathrm{d}t} = -\frac{\partial \mathcal{H}}{\partial q} \quad \text{and} \quad \frac{\mathrm{d}q}{\mathrm{d}t} = \frac{\partial \mathcal{H}}{\partial \mathcal{P}_i} \ .$$

In this case,

$$\frac{\mathrm{d}F}{\mathrm{d}t} = \frac{\partial F}{\partial t} + \{F, \mathcal{H}\}$$

where are called Poisson Brackets; that is, the Poisson Brackets are:

$$\{F, \mathcal{H}\} = \frac{\partial F}{\partial q}\frac{\partial \mathcal{H}}{\partial \mathcal{P}_i} - \frac{\partial F}{\partial \mathcal{P}_i}\frac{\partial \mathcal{H}}{\partial q} \ .$$

The previous equation may be seen as an operator O acting on F, that is

$$OF = \frac{\mathrm{d}F}{\mathrm{d}t}$$

where

$$O = \frac{\mathrm{d}}{\mathrm{d}t} = \frac{\partial}{\partial t} + \{\cdot, \mathcal{H}\}$$

The Poisson Brackets have the following properties:

1. Anticommutative:

$$\{F, G\} = -\{G, F\}$$

2. Distributive:

$$\{F+G,J\}=\{F,J\}+\{G,J\}$$

3. Product:

$$\{FG,J\}=F\{G,J\}+\{F,J\}G$$

4. Jacobi Identity:

$$\{F,\{G,J\}\}+\{J,\{F,G\}\}+\{G\{J,F\}\}=0$$

Mathematicians That Contributed to Analytic Mechanics

- Pierre de Fermat (1601?-1665, French)
- Pierre-Louis Moreau de Maupertuis (1698-1759, French)
- Leonhard Euler (1707-1783, Swiss)
- Jean le Rond d'Alembert (1717-1783, French)
- Giuseppe Lagrangia ((Joseph Lagrange), 1736-1813, Italian)
- Simeon Denis Poisson (1781-1840, French)
- Carl Gustav Jacob Jacobi (1804-1851, German)
- William Rowan Hamilton (1805-1865, English)

0.8 Lecture 8

0.8.1 Matrices, Operators, and Elementary Group [redacted] Theory

Any operator acting on a vector $\vec{x}$ of the space L will create a mapping in the same space between the vector $\vec{x}$ and the vector $\vec{y}$ such that

$$O\vec{x} = \vec{y}\ .$$

If the mapping is one to one, then the operator O may have an inverse such that

$$O^{-1}\vec{y} = \vec{x}$$

or

$$OO^{-1} = O^{-1}O = I$$

where I is the unit operator.

The operator O is said to be linear if

$$O(\vec{x} + \vec{z}) = O\vec{x} + O\vec{z}$$

and

$$O(a\vec{x}) = aO\vec{x}$$

where a is any number.

If the vectors $\vec{x}$ and $\vec{y}$ are given by their components in a frame that expands the space L, we may represent the operator O by a matrix such that the columns representing $\vec{x}$ and $\vec{y}$ are related by

$$\begin{pmatrix} y^1 \\ y^2 \\ \vdots \end{pmatrix} = \begin{pmatrix} O^1_1 & O^2_1 & \cdots \\ O^1_2 & \cdots & \cdots \\ \vdots & & \end{pmatrix} \begin{pmatrix} x^1 \\ x^2 \\ \vdots \end{pmatrix}$$

or

$$y^i = O^i_j x^j\ .$$

The necessary and sufficient condition for the matrix O^i_j to have an inverse is that the Jacobian of the previous equation be different from zero.

If we change the basis from $(\vec{e_1}, \vec{e_2}, \cdots)$ to $(\vec{e_1}', \vec{e_2}', \cdots)$ by a linear transformation A, then the vectors $\vec{x}$ and $\vec{y}$ change as

$$\vec{x}' = A\vec{x}$$

and

$$\vec{y}' = A\vec{y}$$

where A is the matrix for the linear transformation. Now,

$$O\vec{x} = \vec{y}$$

so

$$AO\vec{x} = A\vec{y} = \vec{y}' \ .$$

Introducing between O and $\vec{x}$ the unit matrix $A^{-1}A$,

$$AOA^{-1}A\vec{x} = y'$$

or

$$O'x' = y'$$

where

$$O' = AOA^{-1}$$

which is the equivalent operator of O in the new space described by the new set of coordinates.

It is important to see that (a) operating on a vector $\vec{x}$ of the space (which is written as a single column matrix) is equivalent to $\vec{x}$ operating as a single row matrix (which is the transpose of the column) on the transposed matrix $(\tilde{a})$ where $(\tilde{a}_{ij} = a_{ji})$.

Example in two-dimensions:

$$\begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} a_{11}x_1 + a_{12}x_2 \\ a_{21}x_1 + a_{22}x_2 \end{pmatrix}$$

is equal to

$$\begin{pmatrix} x_1 & x_2 \end{pmatrix} \begin{pmatrix} a_{11} & a_{21} \\ a_{12} & a_{22} \end{pmatrix} = \begin{pmatrix} x_1 a_{11} + x_2 a_{12} \\ x_1 a_{21} + x_2 a_{22} \end{pmatrix}$$

and thus

$$\begin{pmatrix} x_1 & x_2 \end{pmatrix} = \widetilde{\begin{pmatrix} x_1 \\ x_2 \end{pmatrix}}$$

In general, $(\widetilde{ab}) = (\tilde{b})(\tilde{a})$. That is, the transpose of the product of two matrices is equal to the product of the transpose of each matrix in reverse order.

The Adjoint Matrix (a^+)

By definition, a^+ is the transpose matrix with complex conjugate elements. That is,

$$a_{ij}^+ = \tilde{a}_{ij}^* = a_{ji}^*$$

where * indicates the complex conjugate.

The following are properties of the adjoint operation:

Let us (α) and (β) be two matrices or operators; then

$$(\alpha\beta)^+ = \beta^+\alpha^+ \ .$$

Let us a see this with a simple example, α and β both being 2×2 matrices:

$$\alpha = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} \quad , \quad \beta = \begin{pmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{pmatrix}$$

Then

$$\alpha^+ = \begin{pmatrix} a_{11}^* & a_{21}^* \\ a_{12}^* & a_{22}^* \end{pmatrix} \quad \text{and} \quad \beta^+ = \begin{pmatrix} b_{11}^* & b_{21}^* \\ b_{12}^* & b_{22}^* \end{pmatrix}$$

So

$$\alpha\beta = \begin{pmatrix} a_{11}b_{11} + a_{12}b_{21} & a_{11}b_{12} + a_{12}b_{22} \\ a_{21}b_{11} + a_{22}b_{21} & a_{21}b_{12} + a_{22}b_{22} \end{pmatrix}$$

and

$$\beta^+\alpha^+ = \begin{pmatrix} b_{11}^*a_{11}^* + b_{21}^*a_{12}^* & b_{11}^*a_{21}^* + b_{21}^*a_{22}^* \\ b_{12}^*a_{11}^* + b_{22}^*a_{12}^* & b_{21}^*a_{21}^* + b_{22}^*a_{22}^* \end{pmatrix}$$

which is equal to $(\alpha\beta)^+$.

The Hermitian matrix is one that is equal to its adjoint. Thus, a Hermitian matrix (α or β) or operator is such that $\alpha^+ = \alpha$ and $\beta^+ = \beta$, so that $(\alpha\beta)^+ = \beta\alpha$. The anticommutator and the commutator are:

$$(\alpha\beta + \beta\alpha) = (\alpha\beta) + (\alpha\beta)^+ = \text{real matrix operator}$$

and

$$(\alpha\beta - \beta\alpha) = \alpha\beta - (\alpha\beta)^+ = \text{pure imaginary matrix operator}$$

As an example, let us take two Pauli matrices:

$$\alpha = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \quad \text{and} \quad \beta = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}$$

Then the commutator is

$$(\alpha\beta - \beta\alpha) = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}\begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} - \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} = -2i\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$ [redacted]

while the anticommutator is

$$(\alpha\beta + \beta\alpha) = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}\begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} + \begin{pmatrix} 0 & -1 \\ i & 0 \end{pmatrix}\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} = i\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} + i\begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix}$$ [redacted]

$$= \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}$$

A unit matrix or operator leaves the scalar product invariant. By definition, the matrix (a) is unitary if $(a)(a^+) = I$.

When we deal with complex spaces where vectors themselves may be complex functions, like in Quantum Mechanics, the scalar product of two vectors is defined as follows:

$$(\vec{x}, \vec{y}) = (\vec{y}, \vec{x})^*$$

where * indicates the complex conjugate. The product itself is a complex number. Like in any vector space, the following properties must hold true:

$$(\vec{x}, c\vec{y}) = c(\vec{x}, \vec{y}) \quad \text{and} \quad (c\vec{x}, \vec{y}) = c^*(\vec{x}, \vec{y})$$

and

$$(\vec{x}_1 + \vec{x}_2, \vec{y}) = (\vec{x}_1, \vec{y}) + (\vec{x}_2, \vec{y}) \ .$$

The number c is, in general, a complex one.

Next, we will show that if A is a matrix operating on $\vec{x}$, then

$$(A\vec{x}, \vec{y}) = (\vec{x}, A^{+}\vec{y}) \ .$$

First, we expand $\vec{x}$ and $\vec{y}$ on the space basis $\vec{e}_1, \vec{e}_2, \cdots$ Then by definition of the scalar product, the first factor is taken at its complex conjugate. Thus, in two dimensions,

$$(\vec{x}, \vec{y}) = x^{*i}\vec{e}_i \cdot y^j \vec{e}_j = x^{*i} y^j (\vec{e}_i \cdot \vec{e}_j) = x^{*i} y^j g_{ij} \ .$$

The matrix g_{ij} is $(\vec{e}_i, \vec{e}_j)$.

If the basis $\vec{e}_i$ were orthogonal, then $g_{ij} = \delta_{ij} =$ the Kronecker Symbols. [redacted]

For simplicity, let us work in two dimensions and in an orthogonal basis.

$$(A\vec{x}, \vec{y}) = \left[\begin{pmatrix} a_{11}x_1 + a_{12}x_2 \\ a_{21}x_1 + a_{22}x_2 \end{pmatrix}, \begin{pmatrix} y_1 \\ y_2 \end{pmatrix} \right]$$

In the standard scalar product, this is a real vector multiplied by a column vector:

$$\left[\begin{pmatrix} a_{11}^* x_1^* + a_{12}^* x_2^* & a_{21}^* x_1^* + a_{22}^* x_2^* \end{pmatrix} \right] \cdot \begin{pmatrix} y_1 \\ y_2 \end{pmatrix}$$

This is the same as

$$\left[\begin{pmatrix} x_1 \\ x_2 \end{pmatrix}, \begin{pmatrix} a_{11}^* & a_{21}^* \\ a_{12}^* & a_{22}^* \end{pmatrix} \begin{pmatrix} y_1 \\ y_2 \end{pmatrix} \right]$$

or

$$(\vec{x}, A^{+}\vec{y}) \ .$$

The generalization of this simple case to higher dimensions is straightforward. [redacted] As we defined previously, the adjoint matrix of A is A^{+} (the conjugated transport). Now, if A is a unitary matrix which changes basis from $\vec{e}_1, \vec{e}_2, \ldots$ to $\vec{e}\,'_1, \vec{e}\,'_2, \ldots$, we have the following equality:

$$(A\vec{x}, A\vec{y}) = (\vec{x}, A^{+}A\vec{y}) = (\vec{x}, \vec{y}) \ .$$

$A^{+}A = I$ is the unit matrix.

Examples of Linear Operators

1. For any matrix O^i_j:

$$O(\vec{x} + \lambda\vec{y}) = O\vec{x} + \lambda O\vec{y}$$

2. If $L = \int_a^b l(t)\ \mathrm{d}t$, then

$$L\vec{x}(t) = \int_a^b l(t)\vec{x}(t)\ \mathrm{d}t$$

3. If $L = \vec{v}$, then

$$L(\vec{x}) = \vec{v} \cdot \vec{x}$$

Some Geometric Uses of Matrices

Given two vectors, $\vec{u}$ and $\vec{v}$, the area of the parallelepiped formed by them is the modulo or magnitude of of the vector product $\vec{u} \times \vec{v} = |\vec{u}||\vec{v}| \sin\theta$. It can be expressed as a matrix:

$$\vec{u} \times \vec{v} = \begin{vmatrix} \vec{e}_1 & \vec{e}_2 & \vec{e}_3 \\ u_1 & u_2 & u_3 \\ v_1 & v_2 & v_3 \end{vmatrix}$$

where $\vec{e}_i$ is the unit vector along the axis $i(1,2,3)$. If we have three vectors, including $\vec{w}$ along with $\vec{v}$ and $\vec{u}$, then the volume of the parallelepiped built by the three vectors is the base ($|\vec{u} \times \vec{v}|$) multiplied by the projection of $\vec{w}$ on $(\vec{u} \times \vec{v})$: $\vec{w} \cdot (\vec{u} \times \vec{v})$. This is equivalent to substituting $\vec{e}_1, \vec{e}_2, \vec{e}_3$ by $\vec{w}_1, \vec{w}_2, \vec{w}_3$. Thus, the volume is

$$V = \begin{vmatrix} w_1 & w_2 & w_3 \\ u_1 & u_2 & u_3 \\ v_1 & v_2 & v_3 \end{vmatrix}$$

If we square this expression (the determinant of $V =$ the determinant of $\tilde{V}$(the transpose

$$V^2 = \left| \begin{pmatrix} w_1 & w_2 & w_3 \\ u_1 & u_2 & u_3 \\ v_1 & v_2 & v_3 \end{pmatrix} \begin{pmatrix} w_1 & u_1 & v_1 \\ w_2 & u_2 & v_2 \\ w_3 & u_3 & v_3 \end{pmatrix} \right| = \begin{vmatrix} \vec{w} \cdot \vec{w} & \vec{w} \cdot \vec{u} & \vec{w} \cdot \vec{v} \\ \vec{u} \cdot \vec{w} & \vec{u} \cdot \vec{u} & \vec{u} \cdot \vec{v} \\ \vec{v} \cdot \vec{w} & \vec{v} \cdot \vec{u} & \vec{v} \cdot \vec{v} \end{vmatrix}$$

Substituting $\vec{w}$, $\vec{u}$, and $\vec{v}$ by the unitary basis vectors $\vec{e}_1$, $\vec{e}_2$, and $\vec{e}_3$,

$$V^2 = \begin{vmatrix} g_{11} & g_{12} & g_{13} \\ g_{21} & g_{22} & g_{23} \\ g_{31} & g_{32} & g_{33} \end{vmatrix} = g$$

Thus, the square of the volumen limited by $\vec{e}_1$, $\vec{e}_2$, and $\vec{e}_3$ is equal to the determinant of the metric tensor.

In a change of basis,

$$\vec{e}'_i = A^j_i \vec{e}_j$$

and

$$g'_{mn} - A^i_m A^j_n g_{ij} \ .$$

So the determinant of $g'_{mn} = g'$, the expression for which is

$$g' = |A^i_m||A^j_n|g = \Delta^2 g$$

and

$$\frac{1}{\sqrt{g'}} = \Delta \frac{1}{\sqrt{g}} \ .$$

Group Theory: A Review

The operators A, B, and C are said to form a group if

- AB is another operator of the same group
- A^{-1} exists such that $AA^{-1} = A^{-1}A = I$, where A is any element of the group and I is the identity element.
- Or $ABC = (AB)C = A(BC)$ (The Associative Property).

The order of the group is the number m of elements in the group such that for any element A, $A^m = I$ is the identity matrix. Any matrix or operator S that maps the space L into L' and changes the elements of the group into another called an *isomorphism* and relates the elements in the following way:

$$A' = SAS^{-1} \quad , \quad B' = SBS^{-1} \quad , \quad \ldots$$

If $A, B, C, \ldots$ can be mapped to a group of matrices $M(A), M(B), \ldots$ [redacted] in the space L of n dimensions, we say that these groups of matrices are the representation of the abstract group $A, B, C, \ldots$ By changing the basis of the space L, the form of the matrices will change, but there are some parts of them that remain invariant, such as:

1. The sum of the diagonal elements of each matrix, which is called the *character* of the group
2. The determinant of the matrices

The *classes* of a group are the union of all the elements related to each other by a transformation created by an element of the group, say U. Thus, $A_1, A_2, \ldots$ form a class if

$$A_i = UA_jU^{-1} \ .$$

The following are various definitions associated with Group Theory:

1. **Subgroup**: If $H < G$, and H is a group itself under the same operation as G, then H is a subgroup.
2. **Conjugate Subgroup**: equals aHa^{-1} where a is any element of G.
3. **Co-Set of a Subgroup**: equals aH (left co-set) or Ha (right co-set).
4. **Invariant Subgroup**: also called the Self-Conjugate or Normal Divisor. H is invariant if $aHa^{-1} = H$ for any a of the G group. In other words, the left co-set of H equals the right co-set, or $aH = Ha$. That is, H commutes with any element of the G group.
5. **Factor or Quotient Group**: built by the co-sets of n invariant group.
6. **Simple Group**: a group which has no invariant subgroup.
7. **Semi-Simple Group**: a group that has invariant subgroups, none of which are Abelian.

When the representation of the group G is of the following type,

$$D = \left(\begin{array}{cc|c} a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \\ 0 & 0 & c_3 \end{array}\right)$$

then when applying D to a vector (x, y, z), we get

$$(D)\begin{pmatrix} x \\ y \\ 0 \end{pmatrix} = \begin{pmatrix} x' \\ y' \\ 0 \end{pmatrix}$$

and

$$(D)\begin{pmatrix} 0 \\ 0 \\ z \end{pmatrix} = \begin{pmatrix} x' \\ y' \\ z' \end{pmatrix}$$

We note that the component (x, y) of the vector transforms along the same components. In other words, $D(G)$ leaves the subspace (x, y) invariant. However, the space perpendicular to the plane (x, y) has a z-component that, under the transformation, is mixed with the rest of the space such that

$$(0, 0, z) \to (x', y', z')$$

while

$$(x, y, 0) \to (x', y', 0) \ .$$

If, however, $D(G)$ is such that

$$D(G_i) = \left(\begin{array}{cc|c} a_i & b_i & 0 \\ c_i & d_i & 0 \\ \hline 0 & 0 & f_i \end{array}\right)$$

then both spaces are invariant. $(x, y) \to (x', y')$, $z \to z'$, and $D(G_i)$ is thus a reducible representation of G, which implies invariance in some spaces.

If for any vector (x, y, z, u, v) the matrices D are such that the vectors transforms into a vector (x', y', z', u', v') then D is a reducible representation and can be written as

$$D = \left(\begin{array}{ccc|cc} a_{11} & a_{12} & a_{13} & 0 & 0 \\ a_{21} & a_{22} & a_{23} & 0 & 0 \\ a_{31} & a_{32} & a_{33} & 0 & 0 \\ \hline 0 & 0 & 0 & b_{11} & b_{12} \\ 0 & 0 & 0 & b_{21} & b_{22} \end{array}\right)$$

Thus,

$$D = D(a) + D(b)$$

and the space L is also separated as a sum of spaces,

$$L = L(a) + L(b) \ .$$

When D becomes a unitary matrix, we have reached full reducibility.

We will now show that a unitary matrix cannot be anything but completely reducible. Assume, however, that this is not so, and let us work in three dimensions. Thus, if D is assumed to not be completely reducible, then it looks like

$$D = \left(\begin{array}{cc|c} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ 0 & 0 & a_{33} \end{array} \right)$$

But if D is unitary, then its rows and columns are orthogonal to each other. So, the product of row one and row three must be zero. This implies that

$$a_{13} \cdot a_{33} = 0$$

and assuming that $a_{33} \neq 0$, this therefore means that $a_{13} = 0$. The same principle holds for row two and row three, making $a_{23} = 0$. Thus, D must look like

$$\left(\begin{array}{cc|c} a_{11} & a_{12} & 0 \\ a_{21} & a_{23} & 0 \\ \hline 0 & 0 & a_{33} \end{array} \right)$$

which is reducible.

For finite groups, there is always a unitary representation of the group (the proof for this is omitted).

An *Abelian Group* is one where all the elements commute with each other. Thus, if D^i are the components of the group (or their matrix representations) then

$$D^1 D^2 = D^2 D^1 \ .$$

If the eigenvalue of D^i is λ_i and its eigenvector $\vec{x}$, then

$$D^i = \lambda_i \vec{x} \ .$$

Since D^i commutes with any other D of the representation,

$$D D^i \vec{x} = D \lambda_i \vec{x} = \lambda_i D \vec{x}$$

but since $D D^i \vec{x} = D^i D \vec{x}$,

$$\lambda_i (D \vec{x}) = D^i (D \vec{x})$$

which makes D^i, any element of the Abelian Group, just a number λ_i. This means that the dimension of an Abelian Matrix is one (just a number).

In general, if the representation of the group is an $n \times n$ matrix, we say that the *dimension* of the group is n. As we stated before, the *order* of a group is m if for any element of the group a

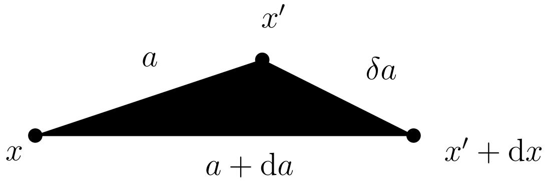
$$a \cdot a \cdot a \cdot \ldots = a^m = I$$

where m is supposed to be the least integer. *Example:* For the rotational group $a = e^{\frac{i\pi}{2}}$, $m = 4$.

0.8.2 Review of Lie Algebra

Lie Algebra was developed by Sophus Lie (1842-1899), a native of Oslo, Norway.

Infinitesimal Transformations



Let us take the following example: one parameter, a, for the transformation that takes $x \to x' = f(x, a)$. The transformation is uniquely determined by the parameter a.

We can go from $x \to x' + \mathrm{d}x'$ by 1) $f(x, a + \mathrm{d}a)$ or 2) by another transformation that starts at x' and whose parameter is b, i.e., $x' \xrightarrow{b} x' + \mathrm{d}x'$:

$$x' \to x' + \mathrm{d}x' \quad \text{by} \quad f(x', b)$$

or

$$x' + \mathrm{d}x' = f(x', b) \ .$$

Since this is equal to two successive transformations (x to x' and x' to $x' + \mathrm{d}x'$), the new parameter b must be a continuous function of a since the transformations are assumed to follow a continuous path. Let is assume that b is a small variation, say some δa. Then, $b \approx \delta a$ may be expressed as a function of a,

$$b \approx \delta a = x(a)\mathrm{d}a \ . \tag{0.8.1}$$

By using a Taylor Expansion of $f(x', a)$ with variation $b \approx \delta a$,

$$x' + \mathrm{d}x' = f(x', \delta a) \approx f(x', 0) + \left(\frac{\partial f(x', a)}{\partial a}\right)_{a=0} \delta a = x' + u(x')\delta a \ .$$

We have used the fact that $f(x',0) \approx x'$ and set $(\frac{\partial f(x',a)}{\partial a})_{a=0}$ equal to $u(x')$, i.e.,

$$u(x') = \Big(\frac{\partial f(x'a)}{\partial a}\Big)_{a=0}$$

and

$$\mathrm{d}x' = u(x')\delta a \ . \tag{0.8.2}$$

From (0.8.1) and (0.8.2),

$$\frac{\mathrm{d}x'}{u(x')} = x(a)\mathrm{d}a \ . \tag{0.8.3}$$

Thus we have separated and related the differential variation of the ”space” (x) which is being transformed and the element of the group that is specified by the parameter a. Next, we will see how a function $F(x)$ changes as $x \to x + \mathrm{d}x$:

$$\mathrm{d}F = \frac{\partial F}{\partial x}\mathrm{d}x$$

but from (0.8.2)

$$\mathrm{d}x = u(x)\delta a$$

so

$$\mathrm{d}F = \Big(u(x)\delta a\frac{\partial}{\partial x}\Big)F \ .$$

In this way, we introduce the operator $X = u(x)\dfrac{\partial}{\partial x}$ so that

$$\mathrm{d}F = \delta a X F \ .$$

$X = u(x)\dfrac{\partial}{\partial x}$ is called the *differential operator of the transformation* $x' = f(x,a)$ that allows one to determine the differential variation of any function $F(x)$ as the original transformation f changes due to a variation of its parameter a. The transformation f forms a group, each element of the group characterized by a parameter a. The corresponding differential operator X also forms a group where the operation among them is the commutator. They are called *generators of the transformation.*

If $F = x'$, then

$$x' = (1 + X\delta a)X \ .$$

Another example is the two-parameter transformation:

$$x' = ax + b \ .$$

The identity is $a = 1, b = 0$. The infinitesimal transformation is represented by the parameters δa and δb, so $\mathrm{d}x = x' - x$ is

$$\mathrm{d}x = x\delta a + \delta b$$

As we have seen, the function $u(x)$ in the expression of x is the coefficient of the increment of the parameter, so

$$u_a(x) = x$$

and

$$u_b(x) = 1 \ .$$

The corresponding infinitesimal operators are $x = u\dfrac{\partial}{\partial x}$:

$$X_a = x\frac{\partial}{\partial x}$$

and

$$X_b = \frac{\partial}{\partial x} \ .$$

The variation of a function F due to the variation of both parameters is

$$\mathrm{d}F = (\delta_a X_a + \delta_b X_b)F \ .$$

The commutators of X_a and X_b are

$$[X_a, X_a] = [X_b, X_b] = 0$$

and

$$[X_a, X_b] = -\frac{\partial}{\partial x} = -X_b \ .$$

Thus, the commutator brings back another infinitesimal operator. That is why the infinitesimal operators form a group under the operation/commutation.

Another example is the one parameter rotational group of transformation in two-dimensions:

$$x' = x\cos\Phi - y\sin\Phi$$

$$y' = x\sin\Phi + y\cos\Phi \ .$$

For the limit $\Phi \to \delta\Phi$, expansing the functions of Φ gives:

$$x' \approx x(1 + \left(\frac{\partial\cos\Phi}{\partial\Phi}\right)_{\Phi=0}\delta\Phi - y(0 + \left(\frac{\partial\sin\Phi}{\partial\Phi}\right)_{\Phi=0}\delta\Phi)) = x - y\delta\Phi \ .$$

$$y' \approx x\delta\Phi + y \ .$$

So,

$$\mathrm{d}x = x' - x = -y\delta\Phi \quad \text{or} \quad u(x) = -y$$

$$\mathrm{d}y = y' - y = x\delta\Phi \quad \text{or} \quad u(y) = x$$

and

$$X_x = u(x)\frac{\partial}{\partial x} = -y\frac{\partial}{\partial x}$$

$$X_y = u(y)\frac{\partial}{\partial y} = x\frac{\partial}{\partial y} \ .$$

The sum of both operators is

$$X = x\frac{\partial}{\partial y} - y\frac{\partial}{\partial x}$$

which is the angular momentum operator around the z-axis. Also,

$$\mathrm{d}F(x,y) = \frac{\partial F}{\partial x}\mathrm{d}x + \frac{\partial F}{\partial y}\mathrm{d}y = \left[- y\delta\Phi\frac{\partial}{\partial x} + x\delta\Phi\frac{\partial}{\partial y}\right]F = (\delta\Phi)XF \ .$$

In the last equation, we have substituted dx and dy from the above expression.

If $F = x^2 + y^2$, the norm of a vector (x, y), the operation of X acting on F leaves F unchanged (the proof for this is straightforward).

As another example, let us analyze the Lorentz transformation in two dimensions, t and x. This transformation may be expressed by the following matrix expression:

$$\begin{pmatrix} x'_0 \\ x'_1 \end{pmatrix} = \begin{pmatrix} \cosh(\xi) & -\sinh(\xi) \\ -\sinh(\xi) & \cosh(\xi) \end{pmatrix} \begin{pmatrix} x_0 \\ x_1 \end{pmatrix}$$

where $x_0 = ct$, $x_1 = x$, and the hyperbolic functions of the parameter ξ for the case of special relativity is given by

$$\cosh(\xi) = \gamma = \frac{1}{\sqrt{1 - \frac{v^2}{c^2}}}$$

and

$$\sinh(\xi) = \frac{v}{c}\gamma$$

It is easy to see that the transformation leaves $x_0^2 - x_1^2$ invariant.

For three dimensional space we have:

$$X_{23} = z\frac{\partial}{\partial y} - y\frac{\partial}{\partial z} \quad , \quad X_{13} = x\frac{\partial}{\partial z} - z\frac{\partial}{\partial x} \quad , \quad X_{12} = y\frac{\partial}{\partial x} - x\frac{\partial}{\partial y} \ .$$

The commutators of these operators are:

$$[X_{23}, X_{13}] = X_{12} \quad , \quad [X_{13}, X_{12}] = X_{23} \quad , \quad [X_{12}, X_{23}] = X_{13} \ .$$

These operators "X" are sometimes indicated by $-iT$ and the transformation A_{12}, as an example, may be written as

$$A_{12}(\beta) = \begin{pmatrix} \cos\beta & -\sin\beta & 0 \\ \sin\beta & \cos\beta & 0 \\ 0 & 0 & 1 \end{pmatrix} = \exp(-i\beta T_{12}) \ .$$

For a small β ($\delta\beta$):

$$\begin{pmatrix} \mathrm{d}x \\ \mathrm{d}y \end{pmatrix} = \left(\frac{\partial A_{12}}{\partial \beta}\right)_{\beta=0} \delta\beta = \begin{pmatrix} 0 & -1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \delta\beta$$

or

$$\begin{pmatrix} \mathrm{d}x \\ \mathrm{d}y \end{pmatrix} = -iT_{12}\delta\beta = i \begin{pmatrix} 0 & i & 0 \\ -i & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \delta\beta$$

or

$$T_{12} = \begin{pmatrix} 0 & -i & 0 \\ i & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

In the language of group theory, the previous transformation is called an *SU2*: the "S" stands for special (the determinant is equal to 1) and "U" stands for unitary, that is, it leaves the scalar product invariant. The unitary matrices have inverses equal to to adjoint ($a^*_{ij} = a^*_{ji}$) and its rows and columns are orthogonal to each other. The number 2 stands for the fact that they are 2×2 matrices.

The SU2 groups are homomorphic to the three-dimensional real rotation which forms what is called the *O3* group. To see it, let us start with a general complex 2×2 matrix u which belongs to the SU2 representation:

$$u = \begin{pmatrix} a & b \\ -b^* & a^* \end{pmatrix} \quad \text{and} \quad u^{-1} = \begin{pmatrix} a^* & -b \\ b^* & a \end{pmatrix}$$

Since a and b are complex, there are four parameters, but only three are independent because of the fact that the determinant should be equal to 1:

$$aa^* + bb^* = 1 \ .$$

Now let h be one matrix of the SU2 representation; let's expand it as a sum of the Pauli Matrices with coefficients x, y, z:

$$\sigma_x = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \quad , \quad \sigma_y = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} \quad , \quad \sigma_z = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$$

So,

$$h = \begin{pmatrix} h_{11} & h_{12} \\ h_{21} & h_{22} \end{pmatrix} = x\sigma_x + y\sigma_y + z\sigma_z = \begin{pmatrix} z & x - iy \\ x + iy & -z \end{pmatrix}$$

Since the determinant of h is equal to 1, we see that $z^2 + y^2 + x^2 = 1$, which expresses the conservation of the scalar product $\vec{r} \cdot \vec{r}$, where $\vec{r} = (x, y, z)$. The values of x, y, z as a function of h_{ij} are

$$x = \frac{h_{12} + h_{21}}{2} \quad , \quad y = \frac{h_{21} - h_{12}}{2i} \quad , \quad z = h_{11} = -h_{22} \ .$$

Since x, y, z are real, the matrix is Hermitian. Thus, h is given as a function of x, y, z. If we change coordinates $(x, y, z \to x', y', z')$ we generate another h' matrix such that $\vec{r} \cdot \vec{\sigma} = (h')$; then,

$$x' = \frac{h'_{12} + h'_{21}}{2} \quad , \quad y' = \frac{h'_{21} - h'_{12}}{2i} \quad , \quad z' = h'_{11} = -h'_{22} \ .$$

Now we introduce the matrix u to go from h to h', $h' = uhu^{-1}$. Then (x', y', z') will be expressed as a function of (x, y, z), and the parameters a and b of u. The result is a rotational matrix that changes (x, y, z) into (x', y', z'). Thus, from a 2×2 SU2 matrix we generate a 3×3 matrix of the O3 group representation that, of course, preserves the scalar product $\vec{r} \cdot \vec{r}$. Let us assume a simple unit matrix u, diagonal only and a function of one parameter, α. The most general one will be

$$u = \begin{pmatrix} e^{-\frac{i\alpha}{2}} & 0 \\ 0 & e^{\frac{i\alpha}{2}} \end{pmatrix} \quad , \quad u^{-1} = \begin{pmatrix} e^{\frac{i\alpha}{2}} & 0 \\ 0 & e^{-\frac{i\alpha}{2}} \end{pmatrix}$$

with

$$h = \vec{r} \cdot \vec{\sigma} = \begin{pmatrix} z & x - iy \\ x + iy & -z \end{pmatrix} \quad \text{and} \quad h' = \vec{r}' \cdot \vec{\sigma}' = \begin{pmatrix} z' & x' - iy' \\ x' + iy' & -z' \end{pmatrix}$$

Thus,

$$h' = uhu^{-1} = \begin{pmatrix} e^{-\frac{i\alpha}{2}} & 0 \\ 0 & e^{\frac{i\alpha}{2}} \end{pmatrix} \begin{pmatrix} z & x - iy \\ x + iy & -z \end{pmatrix} \begin{pmatrix} e^{\frac{i\alpha}{2}} & 0 \\ 0 & e^{-\frac{i\alpha}{2}} \end{pmatrix}$$

From this,

$$\begin{pmatrix} x' \\ y' \\ z' \end{pmatrix} \begin{pmatrix} \cos\alpha & -\sin\alpha & 0 \\ \sin\alpha & \cos\alpha & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = (R) \begin{pmatrix} x \\ y \\ z \end{pmatrix}$$

So the rotation R is in the (x, y) plane around the z-axis.

Lorentz Generators

By expanding the hyperbolic function of ξ and reducing it to a small variation fo ξ, $\delta\xi$, the Lorentz transformation becomes:

$$\begin{pmatrix} x'_0 \\ x'_1 \end{pmatrix} = \left(\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} + \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} \delta\xi \right) \begin{pmatrix} x_0 \\ x_1 \end{pmatrix}$$

or

$$\begin{pmatrix} \mathrm{d}x'_0 \\ \mathrm{d}x'_1 \end{pmatrix} = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} x_0 \\ x_1 \end{pmatrix} \delta\xi = \begin{pmatrix} u_0(x) \\ u_1(x) \end{pmatrix} \delta\xi$$

The infinitesimal operators are

$$X_0 = -x_1 \frac{\partial}{\partial x_0} = -x \frac{\partial}{\partial t}$$

and

$$X_1 = x_0 \frac{\partial}{\partial x_1} = t \frac{\partial}{\partial x} \ .$$

In general, we can write the differential transformation due to an increment of the parameter ξ as

$$\begin{pmatrix} \mathrm{d}x_0 \\ \mathrm{d}x_1 \end{pmatrix} = (l_{01}) \begin{pmatrix} x_0 \\ x_1 \end{pmatrix} \delta\xi$$

The matrix (l_{01}) is called a *generator* of the transformation group. In our case,

$$l_{01} = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$$

In a case of many variables x and one parameter ξ,

$$\begin{pmatrix} \mathrm{d}x_0 \\ \vdots \\ \mathrm{d}x_n \end{pmatrix} = (l_{ij}) \begin{pmatrix} x_0 \\ \vdots \\ x_n \end{pmatrix} \delta\xi$$

and the matrices l are independent of x and ξ, and again they are called generators of the transformation group.

It is easy to see that the function $u_k(x)$ and the (l_{ij}) generator are related by

$$\begin{pmatrix} u_0(x) \\ \vdots \\ u_n(x) \end{pmatrix} = (l_{ij}) \begin{pmatrix} x_0 \\ \vdots \\ x_n \end{pmatrix}$$

where (l_{ij}) is a generator of the group and the infinitesimal operators are

$$X_i = u_i(x)\frac{\partial}{\partial x_i} = \left[(l_{ij})\begin{pmatrix} x_0 \\ \vdots \\ x_n \end{pmatrix}\right]_i \frac{\partial}{\partial x_i}$$

The commutators of X_i are equal to the commutators of the generators (l_{ij}).

In the previous case, the generator is

$$l_{01} = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$$

but if we work with the rotation in four-dimensional space, there are six generators in the Lorentz group corresponding to the six independent rotations which leave the six independent planes (x_{01}, x_1), (x_{01}, x_2), (x_{01}, x_3), (x_1, x_2), (x_2, x_3) invariant under these rotations, and the main invariant may be written as $(x^2 + y^2 + z^2 - t^2)$.

The generators are such that their sum and product as well as their commutator belong to the same set. These properties characterize what is called *Lie Algebra*. The number of commuting generators is called the *rank* of the algebra. The Lorentz group has a rank equal to 2. In general, its commutator obeys the following equation:

$$[l_{\mu\lambda}, l_{\nu\alpha}] = g_{\mu\nu}l_{\lambda\alpha} + g_{\lambda\alpha}l_{\mu\nu} - g_{\mu\alpha}l_{\lambda\nu} - g_{\lambda\nu}l_{\mu\alpha}$$

where the g's are the metric tensor which for special relativity in rectilinear, orthogonal coordinates are are $g_{\mu\nu} = 1, -1, -1, -1$ for the diagonal elements and zero for non-diagonal elements.

For example, take $\mu = \nu = 0$, $\lambda = 3$, and $\alpha = 2$. This means the μ and ν correspond to the time variable while λ corresponds to z

and α to y. Thus,

$$[l_{\mu\lambda}, l_{\nu\alpha}] = [l_{03}, l_{02}] = g_{32}l_{00} + g_{00}l_{32} - g_{32}l_{02} - g_{30}l_{30} \ .$$

The only term not equal to zero is $g_{00}l_{32} = l_{32} = l_{zy}$. So,

$$[l_{03}, l_{02}] = l_{32} \ .$$

In the notation of infinitesimal operators, the commutator for this transformation is

$$[X_3, X_2] = A_1$$

where

$$X_3 = z\frac{\partial}{\partial t} + t\frac{\partial}{\partial z} \quad , \quad X_2 = y\frac{\partial}{\partial t} + t\frac{\partial}{\partial y} \quad , \quad X_1 = z\frac{\partial}{\partial y} - y\frac{\partial}{\partial z} \ .$$

The transformation $f(x \to x', a)$ may have more than one parameter a and the function $u(x)$ as in the previous case may refer to many variables. So let x represent the set: $x_1, x_2, \ldots, x_n$ and a represent the set $a_1, a_2, \ldots, a_r$. So equation (0.8.3) may be written as

$$\mathrm{d}x_0' = u_{ik}(x')X_{kl}(a)\mathrm{d}a_l \tag{0.8.4}$$

where $i = 1, \ldots, i, \ldots, n$ and l and k are: $1, \ldots, k, \ldots, l, \ldots, r$.

The infinitesimal operators are expressed in their general case as

$$X_\rho = u_{i\rho}\frac{\partial}{\partial x_i} \ .$$

It may be shown that, as in the previous example, their commutators are linear combinations of all infinitesimal operators. So if repeated indices imply summation, then

$$[X_\rho, X_\sigma] = C^k_{\rho\sigma} X_k$$

where the coefficients $C^k_{\rho\sigma}$ are real constants and obviously:

$$C^k_{\sigma\rho} = -C^k_{\rho\sigma} \ . \tag{0.8.5}$$

Using the Jacobi Identity for the commutators of X_ρ, X_σ, and X_τ, we find that

$$C^\mu_{\rho\sigma}C^\nu_{\mu\tau} + C^\mu_{\sigma\tau}C^\nu_{\mu\rho} + C^\mu_{\tau\rho}C^\nu_{\mu\sigma} = 0 \ . \tag{0.8.6}$$

Lie showed that by knowing the C's that fulfill equations (0.8.5) and (0.8.6), we may eventually solve for (0.8.4), which will eventually unfold the transformation of $f(x \to x')$. The C coefficients are called *structure constants* for Lie Algebra.

Let us look at the previous analysis regarding the Lie Algebra coefficients of the Lorentz group in a more general manner. Let us see how a metric tensor of the group comes about in relation to the Lie Algebra coefficients.

The fundamental operation for the group of infinitesimal operators is the commutator. Let A and B be two elements of the group.

One element of the group S may generate the other members of the group by using the commutators of A and S and B and S. Thus, let us call P_AS the commutator:

$$(P_A)S = [A, S] \ .$$

If we use a set of $X_0, \dots, X_\mu, \dots, X_\nu$ to expand A and S:

$$A = a_\mu X_\mu \quad \text{and} \quad S = s_\nu X_\nu$$

then

$$(P_A)S = a_\mu s_\nu [X_\mu, X_\nu] = a_\mu s_\nu C^\alpha_{\mu\nu} X_\alpha \ .$$

With a similar expansion for $B = b_\sigma X_\sigma$, we calculate $[B, [A, S]]$, which may be called $(P_BP_A)S$. Thus,

$$(P_BP_A)S = b_\sigma a_\mu s_\nu C^\alpha_{\mu\nu} [X_\sigma, X_\alpha] = a_\mu b_\sigma s_\nu C^\alpha_{\mu\nu} C^\gamma_{\sigma\alpha} X_\gamma \ .$$

Let us sum over μ, σ, and α, so that me may generate a matrix with indexes of of ν and γ:

$$(P_BP_A)_{\nu\gamma} = a_\mu b_\sigma C^\alpha_{\mu\nu} C^\gamma_{\alpha\sigma} \ .$$

The s_μ coefficient is part of the operator S.

The trace of $(P_BP_A)_{\nu\gamma}$ is $\sum_{\nu\gamma} (P_BP_A)_{\nu\gamma} \delta_{\nu\gamma}$, so

$$\text{Trace}(P_BP_A) = a_\mu b_\sigma g_{\mu\sigma}$$

where

$$g_{\mu\sigma} = C^\alpha_{\mu\nu} C^\nu_{\alpha\sigma}$$

and the sum over α and μ is understood.

This trace resembles the scalar product of two vectors, whose components in the system are a_μ and b_σ, which are precisely the components of the infinitesimal operators A and B in the base $X_0, \dots, X_\mu, \dots, X_\nu, \dots$

As we have seen before, the C's are real and antisymmetric in the interchange of the lower indexes. But from the expression of

$g_{\mu\nu}$, the interchange of μ and σ does not vary the value of $g_{\mu\sigma}$. So, this metric matrix is real, symmetric, and leaves invariant the scalar product:

$$(A, B) = g_{\mu\nu} a_\mu a_\nu \ .$$

If a matrix is real and symmetric, it could be diagonalized by a real transformations.

If the determinant of $g_{\mu\nu}$ is zero, then the scalar product

$$(A, B) = 0 \ .$$

Next, we state some results on the Lie Algebra of a semi-simple group without proof:

1. In order for the Lie Algebra to correspond to a Semi-Simple Group (that is, a group with invariant subgroups, none of which are Abelian), no element of the infinitesimal group should be orthogonal to all other elements of the group, or, in other words, the determinant of the g_{ij} must not be equal to zero. This is a theorem proposed by Elie Cartan (1869-1951). Here it is worth observing that an Abelian Lie Algebra is of no major interest because the commutator of its infinitesimal operators vanishes.

2. For a compact Semi-Simple Lie Algebra, $g_{\mu\nu}$ must be negative by definition. Thus, we may choose $g_{\mu\nu} = -\delta_{\mu\nu}$, where $\delta = 1$ for $\mu = \nu$, and otherwise $\delta = 0$. In this case, the structure constants $C^{\alpha}_{\mu\nu}$ are antisymmetric under any interchange of indexes, and of course $(A, A) < 0$ for every elements of the algebra. Also,

$$\text{Trace}(P_B P_A) = a_\mu b_\sigma \delta_{\mu\sigma} = -a_\mu b_\mu$$

Lorentz Generators and Lie Algebra

Let us review some developments of our previous lecture on relativity. The Lorentz transformation that changes the coordinates (t, x, y, z) to (τ, ξ, η, ζ) is a rotation in the plane (x_0, x_1) where the new coordinates are $x_0 = ct$, $x_1 = x$, $x_2 = y$, and $x_3 = z$. Thus, in matrix form:

$$\begin{pmatrix} x'_0 \\ x'_1 \\ x'_2 \\ x'_3 \end{pmatrix} = \begin{pmatrix} c\tau \\ \xi \\ \eta \\ \zeta \end{pmatrix} = \begin{pmatrix} \gamma & -\gamma\frac{v}{c} & 0 & 0 \\ -\gamma\frac{v}{c} & \gamma & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} ct \\ x \\ y \\ z \end{pmatrix}$$

From this expression, we obtain again:

$$c\tau = ct\gamma - \frac{v}{c}\gamma x$$

or

$$\tau = \gamma(t - \frac{v}{c^2}x)$$

and

$$\xi = -\gamma\frac{v}{c}(ct) + x\gamma = \gamma(x - vt) \ .$$

The inverse rotation is done with the inverse matrix. That is,

$$\begin{pmatrix} x_0 \\ x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} ct \\ x \\ y \\ z \end{pmatrix} = \begin{pmatrix} \gamma & \gamma\frac{v}{c} & 0 & 0 \\ \gamma\frac{v}{c} & \gamma & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} c\tau \\ \xi \\ \eta \\ \zeta \end{pmatrix}$$

Whatever the coordinates, the fundamental metric tensor is the same:

$$g_{ij} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix}$$

This metric guarantees that the velocity of light in systems moving one with respect to another with constant velocity is always the same. That is,

$$\mathrm{d}x^2 = c^2\mathrm{d}t^2 \quad , \quad \mathrm{d}\xi^2 = c^2\mathrm{d}\tau^2 \quad ,\text{etc.}$$

for any relative velocity v. The transformation from (t, x, y, z) coordinates to (τ, ξ, η, ζ) will, however, depend on v and it may be seen as a rotation in the planes (x_0, x_1), (x_0, x_2), (x_0, x_3), (x_1, x_2), (x_1, x_3), (x_2, x_3), or any rotation in the four-dimensional space. This last one will be expressed as a product of the six fundamental rotations by some constant angles.

These rotations form a group, the continuous Lorentz Group with six dimensions. They also form a Lie Group, so any differential rotation is expressed in the six fundamental rotations, called generators, of the Lorentz Group. For instance, the previous rotation, for small angles ϵ, may be expressed by the first two terms of the Taylor expansion

$$a_{01}(\xi) \approx \delta_{ij} + (l_{01})_{ij}\epsilon_{01}$$

where a_{01} is a component of the rotational matrix, and

$$l_{01} = \left[\frac{\partial}{\partial\beta}a_{01}(\beta)\right]_{\beta=0} = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

Knowing l_{01}, we may generate the entire transformation in the plane (x^0, x^1).

The previous transformation from (x_0, x_1) to (x_0', x_1') may be expressed by using hyperbolic trigonometry of an angle β:

$$\begin{pmatrix} c\tau \\ \xi \\ \eta \\ \zeta \end{pmatrix} = \begin{pmatrix} \cosh\beta & \sinh\beta & 0 & 0 \\ \sinh\beta & \cosh\beta & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} ct \\ x \\ y \\ z \end{pmatrix}$$

The inverse transformation corresponds to an angle $-\beta$. We define

$$\tanh\beta = -\frac{v}{c} \quad \text{so} \quad \gamma^2 = \frac{1}{1 - \dfrac{\sinh^2\beta}{\cosh^2\beta}} \ .$$

Since $\cosh^2\beta - \sinh^2\beta = 1$,

$$\frac{\sinh^2\beta}{\cosh^2\beta} = 1 - \frac{1}{\cosh^2\beta} = \frac{v^2}{c^2} \ .$$

Thus,

$$\frac{1}{\cosh^2\beta} = 1 - \frac{v^2}{c^2} = \frac{1}{\gamma^2}$$

or

$$\cosh^2\beta = \gamma^2 \quad \text{and} \quad \sinh^2\beta = \gamma^2 - 1 = \gamma^2\frac{v^2}{c^2} \ .$$

The rotations in the planes (x_0, x_2), (x_0, x_3), (x_1, x_2), x_1, x_3), and (x_2, x_3) are expressed by the following matrices:

$$a_{02} = \begin{pmatrix} \cosh\beta & 0 & \sinh\beta & 0 \\ 0 & 1 & 0 & 0 \\ \sinh\beta & 0 & \cosh\beta & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \rightarrow l_{02} = \begin{pmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

$$a_{03} = \begin{pmatrix} \cosh\beta & 0 & 0 & \sinh\beta \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ \sinh\beta & 0 & 0 & \cosh\beta \end{pmatrix} \rightarrow l_{03} = \begin{pmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{pmatrix}$$

$$a_{12} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & \cosh\beta & \sinh\beta & 0 \\ 0 & -\sinh\beta & \cosh\beta & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \rightarrow l_{12} = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

$$a_{13} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & \cosh\beta & 0 & \sinh\beta \\ 0 & 0 & 0 & 0 \\ 0 & -\sinh\beta & 0 & \cosh\beta \end{pmatrix} \rightarrow l_{13} = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \end{pmatrix}$$

$$a_{23} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & \cosh\beta & \sinh\beta \\ 0 & 0 & -\sinh\beta & \cosh\beta \end{pmatrix} \to l_{23} = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & -1 & 0 \end{pmatrix}$$

It can be shown that the commutator of any two of these generators, say $l_{\mu\nu}$ and $l_{\rho\sigma}$, is given as a function of the other generators and the metric tensor. That is,

$$[l_{\mu\nu}, l_{\rho\sigma}] = g_{\nu\sigma} l_{\mu\rho} + g_{\mu\sigma} l_{\nu\rho} - g_{\nu\rho} l_{\mu\sigma} - g_{\mu\sigma} l_{\nu\rho}$$

where

$$g = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix}$$

The set of the generators is called the Lie Algebra of the Lorentz Group.

0.9 Lecture 9: Fundamentals of Today's Cosmology

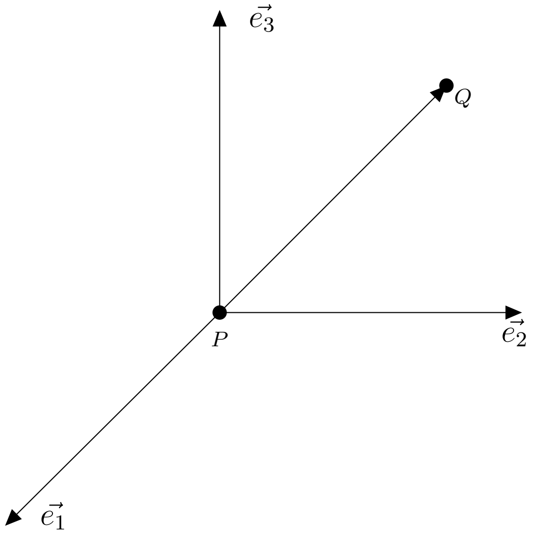
0.9.1 Differential Operators in Curvilinear Coordinates

Call u^1, u^2, u^3 the curvilinear coordinates in three dimensions. $\vec{e_1}, \vec{e_2}, \vec{e_3}$ [redacted] are the "unit vectors" on the three axes, which do not necessarily correspond to unit values. Thus, a vector $\vec{F}$ which begins at the origin P is expressed in curvilinear coordinates as

$$\vec{F} = u^1\vec{e_1} + u^2\vec{e_2} + u^3\vec{e_3}$$

or

$$\mathrm{d}\vec{F} = \mathrm{d}u^1\vec{e_1} + \mathrm{d}u^2\vec{e_2} + \mathrm{d}u^3\vec{e_3} \ .$$

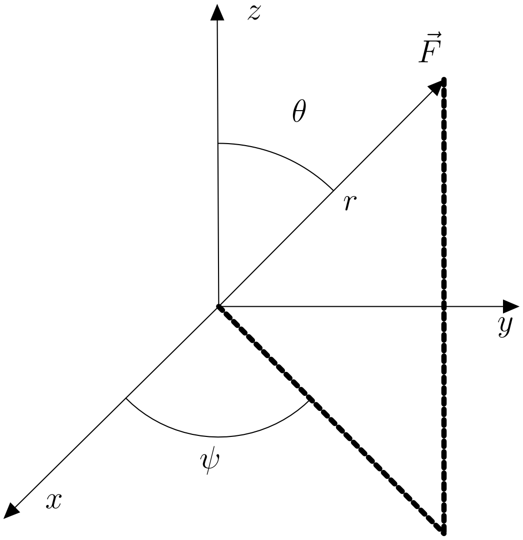


In general, if $\vec{F}$ changes by $\mathrm{d}\vec{F}$, then its extreme Q will describe a curve whose length $\mathrm{d}s$ is such that

$$\mathrm{d}s^2 = g_{ij}\mathrm{d}u^i\mathrm{d}u^j$$

where repeated indexes imply summation. g is the metric tensor, and $g_{ij} = \vec{e_i} \cdot \vec{e_j}$. From now on, we will be working with orthogonal coordinates such that $g_{ij} = 0$, as long as $i \neq j$.

An example in spherical coordinates:



$u^1 = r$, $u^2 = \theta$, and $u^3 = \psi$. With unit vectors of $\hat{r}$, $\hat{\theta}$, and $\hat{\psi}$,

$$\mathrm{d}\vec{F} = \mathrm{d}r\hat{r} + (r\theta\mathrm{d}\theta)\hat{\theta} + (r\sin\theta\mathrm{d}\psi)\hat{\psi}$$

and

$$\vec{e_1} = \hat{r} \quad , \quad \vec{e_2} = (r\theta)\hat{\theta} \quad , \quad \vec{e_3} = (r\sin\theta)\hat{\psi} \ .$$

So, $g_{11} = 1$, $g_{22} = r^2$, and $g_{33} = (r\sin\theta)^2$. Since $\hat{r}$, $\hat{\theta}$, and $\hat{\psi}$ are all orthonormal, all the other components of g are zero, for $i \neq j$. Thus, the g matrix is diagonal, and the determinant is

$$|g| = g_{11}g_{22}g_{33} \ .$$

The element of volume within $\vec{e_1}$, $\vec{e_2}$, and $\vec{e_3}$ is $\mathrm{d}v$,

$$\mathrm{d}v = \sqrt{g}\ \mathrm{d}u^1\mathrm{d}u^2\mathrm{d}u^3 \ .$$

The length of an arch along the coordinates u_i is

$$(\mathrm{d}l)_i = h_i \mathrm{d}u^i$$

where $h_i = \sqrt{g_{ii}}$. Thus, in these spherical coordinates

$$\mathrm{d}\vec{r} = \vec{e_1}\mathrm{d}r + \vec{e_2}\mathrm{d}\theta + \vec{e_3}\mathrm{d}\psi$$

and

$$\mathrm{d}s^2 = (\vec{e_1} \cdot \vec{e_1})(\mathrm{d}r)^2 + (\vec{e_2} \cdot \vec{e_2})(\mathrm{d}\theta)^2 + (\vec{e_3} \cdot \vec{e_3})(\mathrm{d}\psi)^2$$

or

$$g_{ii} = \vec{e_i} \cdot \vec{e_i} \ .$$

The differential lengths along each axis are

$$\begin{aligned}
\mathrm{d}l_1 &= |\vec{e_1}|\mathrm{d}u^1 \\
\mathrm{d}l_2 &= |\vec{e_2}|\mathrm{d}u^2 \\
\mathrm{d}l_3 &= |\vec{e_3}|\mathrm{d}u^3
\end{aligned}$$

Thus, $h_i = |\vec{e_i}|$. $h_i \mathrm{d}u^i$ is a length, and $h_i = \sqrt{g_{ii}}$ makes each of the terms of $\mathrm{d}\vec{r}$ have the same dimensions (length).

The Gradient ∇ of the Function Φ

$$\nabla\Phi = \frac{\partial\Phi}{h_1\partial u^1}\hat{r} + \frac{\partial\Phi}{h_2\partial u^2}\hat{\theta} + \frac{\partial\Phi}{h_3\partial u^3}\hat{\psi}$$

where $h_1 = 1$, $u^1 = r$, $h_2 = r$, $u^2 = \theta$, $h_3 = r\sin\theta$, and $u^3 = \psi$. So in spherical coordinates,

$$\nabla\Phi = \frac{\partial\Phi}{\partial r}\hat{r} + \frac{\partial\Phi}{r\partial\theta}\hat{\theta} + \frac{\partial\Phi}{r\sin\theta\partial\psi}\hat{\psi}\ .$$

The Rotational $\nabla\times$ of the Vector $\vec{F}$

$$\nabla\times\vec{F} = \begin{pmatrix} \hat{r} & \hat{\theta} & \hat{\psi} \\ \dfrac{\partial}{h_1\partial u^1} & \dfrac{\partial}{h_2\partial u^2} & \dfrac{\partial}{h_3\partial u^3} \\ F_1 & F_2 & F_3 \end{pmatrix}$$

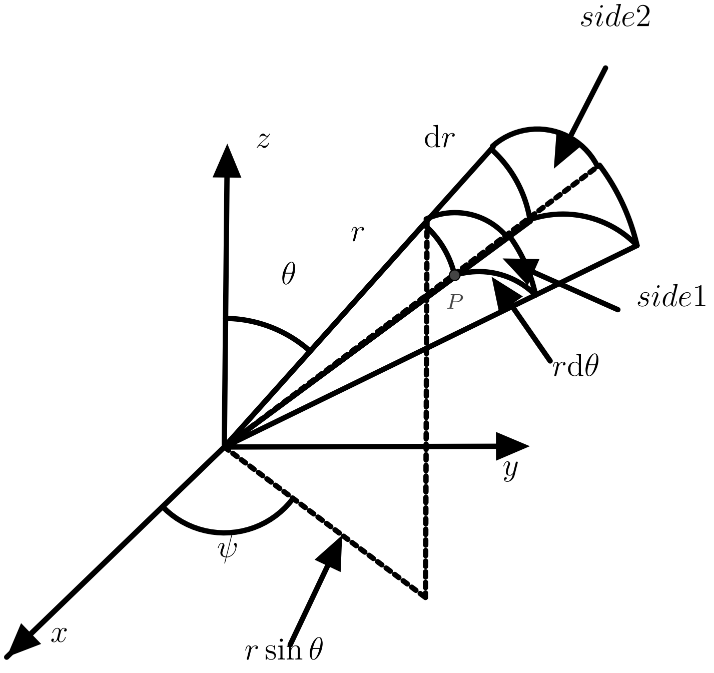
where F_1, F_2, and F_3 are the components of $\vec{F}$ on the three orthogonal directions with unit vectors $\hat{r}$, $\hat{\theta}$, and $\hat{\psi}$. We could resolve this matrix for $\nabla\times\vec{F}$ and find that

$$\nabla\times\vec{F} = \hat{r}\Big(\frac{\partial F_3}{h_2\partial u^2} - \frac{\partial F_2}{h_3\partial u^3}\Big) + \hat{\theta}\Big(\frac{\partial F_1}{h_3\partial u^3} - \frac{\partial F_3}{h_1\partial u^1}\Big) + \hat{\psi}\Big(\frac{\partial F_2}{h_1\partial u^1} - \frac{\partial F_1}{h_2\partial u^2}\Big)$$ [redacted]

where, as before in spherical coordinates, $h_1 = 1$, $h_2 = r$, and $h_3 = r\sin\theta$.

The Divergence $\nabla\cdot$ of the vector $\vec{F}$

From the Gauss Theorem, we may derive the divergence of a vector as the limit of the flux of that vector on an infinitesimal surface, containing the point of interest, per unit volume.



Let us calculate the flux through two opposite sides formed by (side 1) $[r\mathrm{d}\theta, (r\sin\theta)\mathrm{d}\psi]$ and (side 2) $[(r+\mathrm{d}r)\mathrm{d}\theta, r+\mathrm{d}r]$.The unit vector perpendicular to side 1 directed out of the volume is $(\hat{\psi}\times\hat{\theta})$ and for side 2 it is $(\hat{\theta}\times\hat{\psi}) = -(\hat{\psi}\times\hat{\theta})$. The total flux through the two sides is

$$F\cdot(\hat{\theta}\times\hat{\psi})[(r+\mathrm{d}r)^2\sin\theta - r^2\sin\theta]\mathrm{d}\psi\mathrm{d}\theta$$
$$= F\cdot(\hat{\theta}\times\hat{\psi})\sin\theta[(r+\mathrm{d}r)^2 - r^2]\mathrm{d}\psi\mathrm{d}\theta$$

Since $(\hat{\theta}\times\hat{\psi}) = \hat{r}$, and $\vec{F}\cdot\hat{r} = F_r$, we find that the infinitesimal flux through two faces per unit volume is

$$\frac{1}{\mathrm{d}v}\frac{\partial}{\partial r}[F_r(r^2\sin\theta)]\mathrm{d}r\mathrm{d}\theta\mathrm{d}\psi$$

where $\mathrm{d}v = (r^2\sin\theta)\mathrm{d}r\mathrm{d}\theta\mathrm{d}\psi$. Thus, the contribution to the divergence from the two sides is

$$(\nabla \cdot \vec{F})_r = \frac{1}{r^2}\frac{\partial}{\partial r}(F_r r^2) \ .$$

Similar calculations for the other sides reveal

$$\nabla \cdot \vec{F} = \frac{1}{r^2}\frac{\partial}{\partial r}(F_r r^2) + \frac{1}{r\sin\theta}\frac{\partial}{\partial\theta}(F_\theta \sin\theta) + \frac{1}{r\sin\theta}\frac{\partial}{\partial\psi}F_\psi \ .$$

The Laplacian Δ of the Function Φ

Let us take a function $\Phi = \Phi(u^1, u^2, u^3)$. The Laplacian of Φ, $\Delta\Phi$, is the divergence of the gradient of Φ, or $\nabla \cdot (\nabla\Phi)$. Since

$$\vec{F} = \nabla\Phi = \hat{r}\frac{\partial\Phi}{\partial r} + \hat{\theta}\frac{1}{r}\frac{\partial\Phi}{\partial\theta} + \hat{\psi}\frac{1}{r\sin\theta}\frac{\partial\Phi}{\partial\psi}$$

then

$$\nabla\cdot\vec{F} = \Delta\Phi = \frac{1}{r^2}\frac{\partial}{\partial r}\left(\frac{\partial\Phi}{\partial r}r^2\right)+\frac{1}{r\sin\theta}\frac{\partial}{\partial\theta}\left(\frac{1}{r}\frac{\partial\Phi}{\partial\theta}\sin\theta\right)+\frac{1}{r\sin\theta}\frac{\partial}{\partial\psi}\left(\frac{1}{r\sin\theta}\frac{\partial\Phi}{\partial\psi}\right)$$

[redacted]

0.9.2 Potential Theory and the Hamilton-Jacobi Equation

Assume a differential first order linear equation of the following type: we look for the unknown $y(t, q)$ if it comes that

$$F(t, q, p, y) = 0 \tag{0.9.1}$$

where q represents a set of independent variables $q_1, q_2, \ldots$ and p represents a set of independent variables $p_1, p_2, \ldots$. p is related to $y(q, t)$ by

$$p = \frac{\partial y}{\partial q} \ .$$

p and q are called *canonical variables.*

It is possible to solve (0.9.1) by solving two equations called the Hamilton-Jacobi equations. In fact,

$$\frac{\mathrm{d}F}{\mathrm{d}t} = 0 = \frac{\partial F}{\partial t} + \frac{\partial F}{\partial q}\dot{q} + \frac{\partial F}{\partial p}\dot{p} + \frac{\partial F}{\partial y}\frac{\partial y}{\partial q}\dot{q} + \frac{\partial y}{\partial t} \ .$$

By definition, $\dfrac{\partial y}{\partial q} = p$.

In simple cases, F and y do not depend explicitly on t, such that

$$\frac{\partial F}{\partial q}\dot{q} + \frac{\partial F}{\partial p}\dot{p} + \frac{\partial F}{\partial y}p\dot{q} = 0$$

or

$$\Big(\frac{\partial F}{\partial q} + \frac{\partial F}{\partial y}p\Big)\dot{q} + \frac{\partial F}{\partial p}\dot{p} = 0 \ . \tag{0.9.2}$$

If we make

$$\frac{\partial F}{\partial q} + \frac{\partial F}{\partial y}p = -\dot{p} \quad \text{and} \quad \frac{\partial F}{\partial p} = \dot{q} \tag{0.9.3}$$

then we have

$$\dot{p}\dot{q} - \dot{q}\dot{p} = 0 \ .$$

Thus solving the canonical equations (0.9.3) we solve (0.9.2).

If y is not part of F, i.e., if F is independent of y, then (0.9.3) becomes

$$\frac{\partial F}{\partial q} = -\dot{p} \quad \text{and} \quad \frac{\partial F}{\partial p} = \dot{q}\ . \tag{0.9.4}$$

Assume F is the Hamiltonian of a dynamical problem. Let us take the example of a simple harmonic oscillator, where H is the total energy of the system:

$$H = \frac{p^2}{2m} - \frac{\lambda}{2}q^2\ .$$

Directly from this dynamical equation, assuming that the Hamiltonean is constant, we have

$$\frac{\mathrm{d}H}{\mathrm{d}t} = 0 = \dot{p}\frac{p}{m} - \lambda q\dot{q}\ . \tag{0.9.5}$$

Now, $\frac{p}{m} = \dot{q}$ and $\lambda q = \dot{p}$, so (0.9.5) becomes simply

$$\dot{p}\dot{q} - \dot{p}\dot{q} = 0\ .$$

Besides,

$$\frac{\partial H}{\partial t} = \frac{\partial H}{\partial q}\dot{q} + \frac{\partial H}{\partial p}\dot{p} = 0\ .$$

Comparing with (0.9.5),

$$\frac{\partial H}{\partial p} = \frac{p}{m} = \dot{q} \quad \text{and} \quad \frac{\partial H}{\partial q} = -\dot{p} \tag{0.9.6}$$

which agrees with (0.9.4).

Relationship of Potential Theory to the Poisson Bracket

If we have any function $F(p, q, t) = 0$, then

$$\mathrm{d}F = \frac{\partial F}{\partial t} + \left(\frac{\partial F}{\partial q}\dot{q} + \frac{\partial F}{\partial p}\dot{p}\right) = 0$$

and using the Hamilton equations

$$\frac{\partial F}{\partial t} = -\left(\frac{\partial F}{\partial q}\frac{\partial H}{\partial p} - \frac{\partial F}{\partial p}\frac{\partial H}{\partial q}\right) = -\{F, H\}\ .$$

where $\{F, H\}$ is called the Poisson Bracket of F and H.

Differential Equations for Potential Theory

Newton's Law of Gravitation says that the attraction potential between two masses M_1 and M_2 is directly proportional to the product of the masses in indirectly to the distance between them:

$$U = -G\frac{M_1 M_2}{r} \ .$$

Forces are related to their potential by the fact that F is proportional to the gradient of U, ∇U. In three-dimensional Cartesian space,

$$\nabla U = \frac{\partial U}{\partial x}\vec{x} + \frac{\partial U}{\partial y}\vec{y} + \frac{\partial U}{\partial z}\vec{z}$$

where $\vec{x}$, $\vec{y}$, and $\vec{z}$ are the unit vectors along x, y, and z.

The Laplacian of U is ΔU such that

$$\Delta U = \frac{\partial^2 U}{\partial x^2} + \frac{\partial^2 U}{\partial y^2} + \frac{\partial^2 U}{\partial z^2} \ .$$

ΔU may be expressed as the scalar product of the gradient by itself:

$$\Delta = \nabla \cdot \nabla \ .$$

The Laplacian of $\frac{1}{r}$ is zero,

$$\Delta\frac{1}{r} = 0$$

so the Laplacian of the Newton Gravitational Potential is zero.

The gravitational force F between a uniform-density sphere M and a point mass m which are some distance r away from each other in space is

$$F = \nabla U = -G\frac{Mm}{r^2} \ .$$

But if m is inside M, some distance r away from the center of M, then the gravitational force between them is

$$F \propto r \ .$$

Some Typical Potentials in Classic Physics and the Corresponding Differential Equations

The work of any force F when applied from point 1 to point 2 is W such that

$$W = \int_1^2 F(\vec{r}) \cdot \mathrm{d}r = \int_1^2 -\nabla U \cdot \mathrm{d}r = -(U_2 - U_1) \ .$$

Conservative forces are those in which the work W done by them does not depend on the path taken. Obviously if ∇U is integrable, i.e., has a primitive, then this condition is held and the force is conservative.

In addition, a conservative force implies that U is differentiable, that is

$$\frac{\partial^2 U}{\partial x \partial y} = \frac{\partial^2 U}{\partial y \partial x} \ .$$

In fact, by Stoke's Theorem applied to ∇U,

$$\oint \nabla U \cdot \ \mathrm{d}\mathbf{l} = \iint \mathrm{d}s \ [\nabla \times (\nabla U)] \ .$$

The first integral is take on a closed curve which starts in point 1 and ends at the same point. Since this integral is independent of the path taken and it is always zero, then $\nabla \times \nabla U = 0$.

If $\nabla \times \nabla U = 0$, then it means that the following is also zero:

$$\begin{pmatrix} \frac{\partial}{\partial x} & \frac{\partial}{\partial y} & \frac{\partial}{\partial z} \\ \frac{\partial U}{\partial x} & \frac{\partial U}{\partial y} & \frac{\partial U}{\partial z} \end{pmatrix} = \Big(\frac{\partial^2 U}{\partial x \partial y} - \frac{\partial^2 U}{\partial y \partial x} \Big) \vec{i} + \ldots$$

Thus, $\frac{\partial^2 U}{\partial x \partial y}$ must equal $\frac{\partial^2 U}{\partial y \partial x}$, etc. Therefore, a conservative force means a differentiable force or $\nabla \times \nabla U = 0$.

If this is not the case and $\nabla \times \nabla U \neq 0$, the vector $\vec{\xi} = \nabla \times \nabla U$ is generated by the circulation of ∇U around a closed curve.

The Laplacian of some potential functions may or may not be zero, independent of whether they generate conservative or non-conservative forces.

Examples

1.
$$\nabla U = x^2\vec{j} + y^2\vec{i} + 0\vec{k}$$

   Then,
$$\nabla \times \nabla U = \begin{pmatrix} \vec{i} & \vec{j} & \vec{k} \\ \partial x & \partial y & \partial z \\ y^2 & x^2 & 0 \end{pmatrix} = 2\vec{k}(x-y)$$

   which is different from zero unless $x = y$.

   Its Laplacian is zero. In fact,
$$\nabla \cdot \nabla U = \partial_x^2 U + \partial_y^2 U + \partial_z^2 U = 0 \ .$$

   In summary, this example corresponds to a non-conservative force, $\nabla \times \nabla U \neq 0$ with a null Laplacian, $\nabla^2 U = 0$.

2. If instead of specifying the force (i.e., ∇U) we specify U as a simple differentiable potential, we know that if
$$\nabla \times \nabla U = 0$$

   the force ∇U will be conservative, but its Laplacian may be different from zero. For example, take the potential
$$U = x^3 + y^3 \ .$$

   For this, $\nabla \times \nabla U = 0$. However, since $\dfrac{\partial^2 U}{\partial x^2} = 6x$ and $\dfrac{\partial^2 U}{\partial y^2} = 6y$, then
$$\Delta U = 6(x+y)$$

   which is only zero for $x = -y$ or $x = y = 0$.

   Let us take another example,
$$U = (x^2 + y^2 + z^2)^{-\frac{1}{2}} \ .$$

Then,

$$\Delta U = \frac{\partial}{\partial x}[(-x)(x^2+y^2+z^2)^{-\frac{3}{2}}]+\frac{\partial}{\partial y}[(-y)(x^2+y^2+z^2)^{-\frac{3}{2}}]+\frac{\partial}{\partial z}[(-z)(x^2+y^2+z^2)^{-\frac{3}{2}}] = 0$$

[redacted]

and also

$$\nabla \times \nabla U = 0 \ .$$

Thus, the potential is conservative and its Laplacian ΔU is zero.

3. Let us take a Hook-type of potential, where $F \propto r$, or

$$U = \frac{r^2}{2} = \left(\frac{x^2}{2} + \frac{y^2}{2} + \frac{z^2}{2}\right) \ .$$

Thus,

$$\nabla U = x\vec{i} + y\vec{j} + z\vec{k}$$

and

$$\nabla \times \nabla U = 0 \ .$$

However,

$$\Delta U = \left(\frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2}\right)\left(\frac{x^2}{2} + \frac{y^2}{2} + \frac{z^2}{2}\right) = 3$$

Therefore, the Hook force does not hold a potential whose Laplacian equals 0.

If the force is Newtonian, where $F \propto \frac{1}{r^2}$, plus a Hook-type force, say

$$F = -\frac{a}{r^2} + br$$

then the potential will be such that

$$F = -\nabla U \ .$$

And so,

$$U = -\frac{a}{r} - \frac{br^2}{2}$$

and the Laplacian of U will be

$$\Delta U = \frac{1}{r^2}\frac{\partial}{\partial r}\left(\frac{\partial U}{\partial r}r^2\right) = -3b$$

$\Delta U = -3b$ is a Poisson-type equation.

0.9.3 Simplifications Regarding the the Space-time ”Length” Invariant, the Geodesic Equation, the Red Shift, Black Holes, and the Einstein Approximation

Next, we will be looking at different approximations that will be indicated by S1, S2, . . . , etc. and their application to:

1.) The Space-time ”Length” Invariant

$$\mathrm{d}s^2 = g_{\mu\mu}(x^\mu)^2 \ . \tag{0.9.7}$$

S1: Assuming that the prevailing term is $g_{00}(x^0)^2$, then

$$\mathrm{d}s^2 = c^2(\mathrm{d}\tau)^2 = g_{00}(\mathrm{d}x^0)^2 \ . \tag{0.9.8}$$

Since $v^0 = \dfrac{\mathrm{d}x^0}{\mathrm{d}\tau}$,

$$1 = \left(\frac{v^0}{c}\right)^2 g_{00} \ . \tag{0.9.9}$$

S2: Let us use only two coordinates; x^2 and x^3 are neglected. Thus,

$$\mathrm{d}s^2 = g_{00}(\mathrm{d}x^0)^2 + g_{11}(\mathrm{d}x^1)^2$$

or

$$c^2 = g_{00}(v^0)^2 + g_{11}(v^1)^2 \ . \tag{0.9.10}$$

2.) The Geodesic Equation

$$\frac{\mathrm{d}v^\mu}{\mathrm{d}\tau} = -\Gamma^\mu_{\alpha\sigma} v^\alpha v^\sigma \ . \tag{0.9.11}$$

As before, let us use only x^0 and x^1 (neglecting x^2 and x^3); thus μ is 0 or 1. So,

$$\frac{\mathrm{d}v^0}{\mathrm{d}\tau} = -2\Gamma^0_{01} v^0 v^1 = -g^{00} g_{00,1} v^0 v^1 \ . \tag{0.9.12}$$

Now,

$$g_{00,1} = \frac{\partial g_{00}}{\partial x^1} = \frac{\partial g_{00}}{\partial \tau}\frac{\partial \tau}{\partial x^1} = \frac{\partial g_{00}}{\partial \tau}\frac{1}{v_1} \ .$$

Substituting into the previous equation,

$$\frac{\mathrm{d}v^0}{\mathrm{d}\tau} = -\frac{1}{g_{00}}\frac{\partial g_{00}}{\partial \tau}v^0 \ .$$

Integrating this equation, we get

$$\ln v^0 = -\ln g_{00}$$

or

$$v^0 g_{00} = \xi c \tag{0.9.13}$$

where ξ is a constant and c is the speed of light and $v^0 = c\frac{\mathrm{d}t}{\mathrm{d}\tau}$. Observe that for $\xi = g_{00}$, $\frac{\mathrm{d}t}{\mathrm{d}\tau} = 1$. In a flat space-time, where

$$\frac{\mathrm{d}t}{\mathrm{d}\tau} = \gamma = \frac{1}{\sqrt{1-\frac{v^2}{c^2}}}$$

it means that $v \ll c$. So if our body is accelerated with an acceleration $\frac{\mathrm{d}v^0}{\mathrm{d}\tau}$, then ξ may correspond to its initial velocity.

For $\mu = 1$, the geodesic is written as (neglecting coordinates x^2 and x^3)

$$\frac{\mathrm{d}v^1}{\mathrm{d}\tau} = -\Gamma^1_{00}(v^0)^2 - \Gamma^1_{11}(v^1)^2 = (\frac{1}{2}g^{11}g_{00,1})(v^0)^2 - (\frac{1}{2}g^{11}g_{11,1})(v^1)^2$$

which can be simplified to

$$\frac{\mathrm{d}v^1}{\mathrm{d}\tau} = \frac{1}{2}g^{11}(g_{00,1}(v^0)^2 - g_{11,1}(v^1)^2) \ . \tag{0.9.14}$$

To derive the Newton Law, we neglect the second term, assuming it to be much smaller than the first one. Thus,

$$\frac{\mathrm{d}v^1}{\mathrm{d}\tau} \approx \frac{1}{2}g^{11}g_{00,1}(v^0)^2 \ . \tag{0.9.15}$$

But

$$\frac{\mathrm{d}v^1}{\mathrm{d}\tau} = \frac{\mathrm{d}v^1}{\mathrm{d}x^0}\frac{\mathrm{d}x^0}{\mathrm{d}\tau} = \frac{\mathrm{d}v^1}{\mathrm{d}x^0}v^0$$

so substituting into the previous equation,

$$\frac{\mathrm{d}v^1}{\mathrm{d}x^0} = \frac{1}{2}g^{11}g_{00,1}v^0 \ .$$

Now using equation (0.9.9),

$$\frac{\mathrm{d}v^1}{\mathrm{d}x^0} = \frac{1}{2}g^{11}g_{00,1}\frac{c}{\sqrt{g_{00}}} \ .$$

Since $g^{1\nu} = 0$ for $\nu \neq 1$, then using the lower indexes technique

$$g^{11}g_{00,1} \equiv g^{1\nu}g_{00,\nu} = -g_{00,1} \ .$$

The minus sign is to ensure that g^{11} brings a negative sign:

$$\frac{\mathrm{d}v^1}{\mathrm{d}x^0} = -\frac{c}{2}\frac{g_{00,1}}{\sqrt{g_{00}}} = -c(\sqrt{g_{00}})_{,1} \ .$$

Since $\mathrm{d}x^0 = c\mathrm{d}t$, we find that the acceleration along x^1 is

$$\frac{\mathrm{d}v^1}{\mathrm{d}t} = -c^2(\sqrt{g_{00}})_{,1} \ . \qquad (0.9.16)$$

Equation (0.9.16) resembles the Newton equation of motion if $\sqrt{g_{00}}$ is proportional to the gravitational potential function of x^1. Since g_{00} for small curvatures will be close to 1, we can write

$$g_{00} = (1 + 2\Phi)$$

where Φ is a small quantity. Next, using the Taylor expansion we can express the acceleration for a unit mass as

$$\alpha = -(\sqrt{g_{00}})_{,1}c^2 = -c^2(\sqrt{1+2\Phi})_{,1} \approx -c^2(1+\Phi)_{,1} = -c^2\nabla_1\Phi \qquad (0.9.17)$$

using the approximation: $\sqrt{1+2\Phi} \approx 1 + \Phi$.

3.) The Red Shift Due to Gravitational Effects and the Expansion of the Universe

Using the simplification of (0.9.7), we can write

$$(\Delta s)^2 = c^2\tau^2 = g_{00}c^2T^2$$

where τ is the period of the light wave at the point where light is emitted and T is the period observed after the light has been affected by a gravitational potential which is implicit in g_{00}. For a low gravitational effect,

$$\sqrt{g_{00}} \approx \sqrt{1 - 2\frac{U}{c^2}}$$

where U is the gravitational potential $U \approx \frac{6.17}{9}10^{-9}$. Thus $g_{00} < 1$ and $T = \frac{\tau}{\sqrt{g_{00}}} > \tau$. Thus the frequency $\nu = \frac{1}{T} < \frac{1}{\tau}$; then the original light seen at the source is affected by the gravitational field, and the observer far from the source sees the light at a lower frequency (towards red). The parameter used to characterize this phenomena is called z:

$$z = \frac{\nu_E - \nu_0}{\nu_0}$$

where ν_E is the frequency of emission at the source and ν_0 is the observed frequency.

Observe that if $T = t_1 - t_0$ and $\tau = \tau_1 - \tau_0$, and assuming that the light is emitted at τ_1 and received at t_1, then the distance in classical or special relativity between source and observer is $c(t_1 - \tau_1) = r_0$ or $c(t_0 - \tau_0) = r_0$. This is not the case if $g_{00} \neq 1$, as in general relativity.

The red shift can be affected by more than just gravitation: the Doppler effect as well as the expansion of the universe play large roles as well. The expansion of the universe refers to how, as time goes on, the distance between two points in the universe will change as galaxies drift away from each other and as stars within a galaxy travel away from each other. If we assume that the space coordinates x^1, x^2, x^3 change as $a(t)x^1$, $a(t)x^2$, $a(t)x^3$, then we may see that the red shift is proportional to a. In order to see this, we use the geodesic equation for the x^0 coordinates:

$$\frac{\mathrm{d}v^0}{\mathrm{d}\tau} = -\Gamma^0_{\mu\nu} v^\mu v^\nu$$

where $v^\mu = \dfrac{\mathrm{d}x^\mu}{\mathrm{d}\tau}$.

In accordance with the above assumptions, the invariant $\mathrm{d}s^2$ changes as

$$\mathrm{d}s^2 = g_{00}c^2\mathrm{d}t^2 + \delta_{ii}\mathrm{d}x^i\mathrm{d}x^i \to \mathrm{d}s^2 = g_{00}c^2\mathrm{d}t^2 + g_{ii}\mathrm{d}x^i\mathrm{d}x^i$$

Where $g_{ii} = -a^2(t)\delta_{ii}$ and $\delta_{ii} = 1$ for $i = 1, 2, 3$. In other words, the metric tensor $g_{\mu\nu}$ is

$$(g_{\mu\nu}) = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -a^2(t) & 0 & 0 \\ 0 & 0 & -a^2(t) & 0 \\ 0 & 0 & 0 & -a^2(t) \end{pmatrix}$$

Then the only Christoffel symbols different from zero are

$$\Gamma^i_{0i} = \Gamma^i_{i0} = \frac{\dot{a}}{ca} \quad \text{and} \quad \Gamma^0_{ii} = \frac{\dot{a}a}{c}$$

($\dot{a}$ refers to the derivative of a with respect to t). To prove this, let us calculate Γ^0_{11} and Γ^1_{10}.

$$\Gamma^1_{01} = \frac{1}{2}g^{1k}(g_{1k,0} + g_{k0,1} - g_{10,k}) = \frac{1}{2}g^{11}g_{11,0}\ .$$

The only $g^{1k} \neq 0$ is g^{11} and the only $g_{1k} \neq 0$ is g_{11}. Now, $g^{11} = -\dfrac{1}{a^2(t)}$ and $g_{11,0} = \dfrac{\mathrm{d}}{c\mathrm{d}t}(-a^2(t)) = -\dfrac{2}{c}a\dot{a}$. Thus,

$$\Gamma^1_{01} = \frac{\dot{a}}{ca}\ . \tag{0.9.18}$$

In a similar way,

$$\Gamma^0_{11} = \frac{1}{2} g^{00}(g_{10,1} + g_{10,1} - g_{11,0}) = \frac{1}{2} g^{00}(-g_{11,0}) \ .$$

$g^{00} = 1$, so

$$\Gamma^0_{11} = \frac{a\dot{a}}{c} \ . \tag{0.9.19}$$

Therefore, the acceleration of a photon according to the geodesic equation is

$$\mathring{A} = -\Gamma^0_{11} v^i v^i = -\frac{a\dot{a}}{c} v^i v^i \tag{0.9.20}$$

for $i = 1, 2, 3$. On the other hand, from the four-velocity $U = \gamma(c, \vec{u})$,

$$P^\mu = m_0 U = (mc, \vec{p}) = (\frac{E}{c}, \vec{p}) \ . \tag{0.9.21}$$

For photons, the momentum $p = \dfrac{E}{c} = \dfrac{h\nu}{c}$.

The magnitude of P^μ squared is $P^\mu P_\mu = \dfrac{E^2}{c^2} - (\vec{p})^2$, but since $(\vec{p})^2 = \dfrac{E^2}{c^2}$, then $P^\mu P_\mu = 0$, or

$$m^2 c^2 - (\vec{p})^2 = \frac{E^2}{c^2} - p^i p^i = 0 \ . \tag{0.9.22}$$

Since $v^i = \dfrac{\mathrm{d}x^i}{\mathrm{d}\tau} = \gamma \dfrac{\mathrm{d}x^i}{\mathrm{d}t}$ and $m_0 \gamma = m$,

$$m_0 v^i = m \frac{\mathrm{d}x^i}{\mathrm{d}t} = p^i \ .$$

Thus, (0.9.20) may be written as

$$\mathring{A} = -\frac{\dot{a}a}{c m_0^2} p^i p^i$$

and from (0.9.22)

$$\mathring{A} = -\frac{\dot{a}a}{c m_0^2} \frac{E^2}{c^2} \frac{1}{a^2} = -\frac{\dot{a}}{a c^3 m_0^2} E^2 \ . \tag{0.9.23}$$

Also,

$$\overset{\circ}{A} = \frac{\mathrm{d}U}{\mathrm{d}\tau} = \gamma\frac{\mathrm{d}}{\mathrm{d}t}[\gamma(c,\vec{u})] = \frac{m}{m_0}\frac{\mathrm{d}}{\mathrm{d}t}(\frac{m}{m_0}c,\gamma\vec{u}) \ .$$

For a photon,

$$\overset{\circ}{A} = \frac{E}{m_0^2c^3}\frac{\mathrm{d}E}{\mathrm{d}t} \ . \tag{0.9.24}$$

Comparing (0.9.23) and (0.9.24),

$$-\frac{\dot{a}}{a}E^2 = E\frac{\mathrm{d}E}{\mathrm{d}t}$$

or

$$\frac{1}{a}\frac{\mathrm{d}a}{\mathrm{d}t} = -\frac{1}{E}\frac{\mathrm{d}E}{\mathrm{d}t} \to \ln E = \ln(\frac{\beta}{a})$$

or

$$E = \frac{\beta}{a}$$

where β is a constant. Since for photons $E = h\nu$, then as $a(t)$ increases (universal expansion) the frequency of the received light decreases (shifts towards red).

4.) Classic Visualization of Black Holes

Gravitational attraction might be overcome by some high initial velocity. Assuming no attraction at a large distance (∞) from the central attracting body with mass M, then by conservation of energy we have

$$\frac{mv^2}{2} - \frac{GMm}{r} = \frac{1}{2}mv_\infty^2 \ .$$

The minimum escape velocity corresponds to $v_\infty = 0$. Thus, at some distance r from the center of the body M, a body with velocity v_e will escape from M's gravitational field at $v_e \geq \sqrt{2\frac{GM}{r}}$. For $r = R$ (the Earth's radius) and M being the mass of the Earth,

$$v_e = \sqrt{2|U|} = \sqrt{2(6.17\times 10^7)} = 1.16\times 10^4\frac{m}{s} \ .$$

Since the velocity of sound is $v_s = 332\frac{m}{s}$,

$$\frac{v_e}{v_s} = 33.7 \text{ Mach}$$

For a body with very large M or very small r, v_e may reach the velocity of light. Places where the escape velocity is greater than the velocity of light, so that no energy is released in the form of mass or light, are called black holes.

Assuming that $v_e = c$, then

$$\frac{M}{r} = \frac{c^2}{2G} \approx 10^{27}\frac{kg}{m} \; .$$

Thus, if M is the mass of a galaxy ($\approx 10^{41}$ kg), $r \approx 10^{14}$ m. If M is the mass of the sun ($\approx 10^{30}$ kg), $r \approx 1$ km. If M is the mass of the earth ($\approx 10^{27}$ kg), $r \approx 1$ m.

Simplified Black Hole Derivations Using Equations (0.9.10) **and** (0.9.13)

The idea is to find the ratio $\frac{\mathrm{d}t}{\mathrm{d}r}$ or $\frac{c\mathrm{d}t}{\mathrm{d}r} = \frac{v^0}{v^1}$ as a function of the main metric tensor g_{00} which we associate with a potential similar to the Newton one. Since $g_{00} \approx 1 - \frac{2m}{r}$, we will attempt to find $\frac{\mathrm{d}t}{\mathrm{d}r}$ as a function of $1 - \frac{2m}{r}$.

Let us take the equations

$$v^0 g_{00} = \xi c \tag{0.9.25}$$

and

$$c^2 = g_{00}(v^0)^2 + g_{11}(v^1)^2 \; . \tag{0.9.26}$$

Eliminating v^0 between (0.9.25) and (0.9.26), we get

$$c^2 g_{00} = \xi^2 c^2 + g_{11} g_{00}(v^1)^2 \; . \tag{0.9.27}$$

Assuming $g_{11}g_{00}$, which is the determinant of $|g|$, is equal to -1 as it is in the Lorentz transformation and the Schwarzschild tensor matrix, then

$$c^2 g_{00} = \xi^2 c^2 - (v^1)^2$$

or

$$g_{00} = \xi^2 - \left(\frac{v^1}{c}\right)^2 . \tag{0.9.28}$$

Thus,

$$c\frac{\mathrm{d}t}{\mathrm{d}r} = \frac{c\dfrac{\mathrm{d}t}{\mathrm{d}\tau}}{\dfrac{\mathrm{d}r}{\mathrm{d}\tau}} = \frac{v^0}{v^1}$$

where τ is the proper time of the light. For $\frac{v^0}{v^1}$, we can use equations (0.9.25) and (0.9.28), so

$$\frac{v^0}{v^1} = \frac{\xi}{g_{00}\sqrt{\xi^2 - g_{00}}} .$$

Now we assume that $g_{00} = 1 - \frac{2m}{r}$, so

$$\frac{v^0}{v^1} = \frac{\xi}{(1 - \frac{2m}{r})\sqrt{\xi^2 - (1 - \frac{2m}{r})}}$$

or

$$\mathrm{d}t = \frac{\mathrm{d}r\frac{\xi}{c}}{(1 - \frac{2m}{r})\sqrt{\xi^2 - (1 - \frac{2m}{r})}}$$

For $r = 2m$ minus some infinitesimal value ϵ, $r = 2m - \epsilon$ or $2m = r + \epsilon$, then

$$1 - \frac{2m}{r} = 1 - \frac{r+\epsilon}{r} \approx \frac{\epsilon}{r}$$

and

$$\xi^2 - (1 - \frac{2m}{r}) \approx \xi^2 - \frac{\epsilon}{r}$$

Thus,

$$\mathrm{d}t = \frac{\mathrm{d}r}{\frac{c\epsilon}{r}\sqrt{\xi^2 - \frac{\epsilon}{r}}}$$

Neglecting ϵ^2 in the Taylor expansion,

$$\sqrt{\xi^2 - \frac{\epsilon}{r}} \approx \xi^2 + \frac{1}{2}\frac{-\frac{\epsilon}{r}}{\left(\xi^2 - \frac{\epsilon}{r}\right)^{\frac{3}{2}}_{\frac{\epsilon}{r}=0}} + \cdots = \xi^2 - \frac{1}{2}\frac{\epsilon}{r\xi^3}$$

In the denominator of dt we again neglect ϵ^2, so

$$\frac{\epsilon}{r}(\xi^2 - \frac{1}{2}\frac{\epsilon}{r}\frac{1}{\xi^3}) \approx \xi^2\frac{\epsilon}{r}$$

Thus

$$t = \int_0^t \mathrm{d}t = \frac{1}{c}\int_0^r \frac{\mathrm{d}r}{\epsilon}\frac{r}{\xi^2} = \frac{1}{c\epsilon\xi^2}\frac{r^2}{2} \tag{0.9.29}$$

For $r \approx 2m$, $\epsilon \to 0$ and $t \to \infty$. Thus, a particle at this distance r will take an infinite time t to leave the attraction of the black hole.

5.) The Einstein Relationship in Different Approximations

The Einstein assumption for space with gravitational forces is to constrain the curvature tensor in the form given by the Ricci tensor to a value of zero. Thus,

$$R_{\nu\rho} = 0$$

or

$$R_{\nu\rho} = \Gamma^{\sigma}_{\nu\sigma,\rho} - \Gamma^{\sigma}_{\nu\rho,\sigma} + \Gamma^{\alpha}_{\nu\sigma}\Gamma^{\sigma}_{\rho\alpha} - \Gamma^{\alpha}_{\nu\rho}\Gamma^{\sigma}_{\sigma\alpha} = 0\ .$$

Let us now assume that (**ER1**):

1. $g_{\mu\nu}$ is diagonal.

2. $g_{\mu\nu,0} = 0$.

3. $g_{\mu\nu,i} = 0$, for $i = 2, 3$.

Thus, the only $R_{\nu\rho}$ different from zero are R_{00}, R_{11}, R_{22}, and R_{33}.

$$\begin{aligned}
R_{00} &= -\Gamma^1_{00,1} + \Gamma^1_{00}\Gamma^0_{01} - \Gamma^1_{00}\Gamma^1_{11} - \Gamma^1_{00}\Gamma^2_{21} - \Gamma^1_{00}\Gamma^3_{31}\\
R_{11} &= \Gamma^0_{10,1} + \Gamma^2_{12,1} + \Gamma^3_{13,1} + \Gamma^2_{12}\Gamma^2_{12} + \Gamma^3_{13}\Gamma^3_{13} + \Gamma^0_{10}\Gamma^0_{10} - \Gamma^1_{11}\Gamma^0_{01} - \Gamma^1_{11}\Gamma^2_{12} - \Gamma^1_{11}\Gamma^3_{13}\\
R_{22} &= \Gamma^3_{23,2} - \Gamma^1_{22,1} + \Gamma^2_{12}\Gamma^1_{22} + \Gamma^1_{22}\Gamma^2_{12} + \Gamma^3_{23}\Gamma^3_{23} - \Gamma^1_{22}\Gamma^1_{11} - \Gamma^1_{22}\Gamma^0_{01} - \Gamma^1_{22}\Gamma^2_{12} - \Gamma^1_{22}\Gamma^3_{13}\\
R_{33} &= -\Gamma^1_{33,1} - \Gamma^2_{33,2} + \Gamma^3_{31}\Gamma^1_{23} + \Gamma^2_{33}\Gamma^3_{23} - \Gamma^1_{33}\Gamma^0_{01} - \Gamma^1_{33}\Gamma^2_{21} - \Gamma^1_{33}\Gamma^1_{11}
\end{aligned}$$

Subapproximation for the Einstein Relationship

Let us assume the following (**ER2**):

1. Only x^0 is relevant.
2. $v^1, v^2, v^3 \ll c$.

Therefore only g_{00} is considered. Thus, the Christoffel symbols which are relevant are

$$\Gamma^0_{01} = \frac{1}{2} g^{00} g_{00,1}$$

and

$$\Gamma^1_{00} = \frac{1}{2} g^{11}(-g_{00,1}) \ .$$

Thus

$$R_{00} = -\Gamma^1_{00,1} + \Gamma^0_{01}\Gamma^1_{00} = \frac{1}{2}(g^{11}g_{00,1})_{,1} = \frac{1}{4} g^{11} g^{00} (g_{00,1})^2$$

and

$$R_{11} = \Gamma^0_{10,1} + \Gamma^0_{10}\Gamma^0_{10} = \frac{1}{2}(g^{00}g_{00,1})_{,1} + \frac{1}{4}(g^{00}g_{00,1})^2 \ .$$

$R_{00} = 0$ means

$$R_{00} = g_{00,11} - \frac{1}{2} g^{00} (g_{00,1})^2 = 0 \ .$$

$R_{11} = 0$ means

$$R_{11} = g^{00}_{,1} g_{00,1} + g^{00} g_{00,11} + \frac{1}{2}(g^{00})^2 (g_{00,1})^2 = 0 \ .$$

If $S = R_{00}g^{00} + R_{11}$, then

$$S = 2g^{00}g_{00,11} + g^{00}_{,1}g_{00,1} = 0 \ .$$

Now,

$$g^{00}_{,1} = (\frac{1}{g_{00}})' = -\frac{g_{00,1}}{(g_{00})^2}$$

so S becomes

$$S = 2g^{00}g_{00,11} - \frac{(g_{00,1})^2}{(g_{00})^2} = 0 \ .$$

Thus,

$$S = 0 \quad \text{if} \quad g_{00,11} = 0 \tag{0.9.30}$$

which is a type of Laplace equation for the potential g_{00}.

Equation (0.9.30) might be put into scalar form by extending the subscript 1 to $\mu\nu$:

$$g_{00,11} \to g_{00,\mu\nu} \ .$$

Let us introduce a generalized Laplacian $\Box_{\mu\nu} \equiv_{,} \mu\nu$, so that we might write the scalar form of $g^{\mu\nu}g_{00,\mu\nu}$ as

$$g^{\mu\nu}\Box_{\mu\nu}g_{00} = 0 \ . \tag{0.9.31}$$

Simplifications for the Ricci Tensor and the Gravitational Einstein Equations in a New Approximation

Now let us make the following new assumptions (**ER5**):

1. $g_{\mu\nu}$ is diagonal.
2. $g_{00} = -\dfrac{1}{g_{11}}$, which is a function of x^1 only.
3. $g_{22} = g_{33} = -1$.

According to the simplifications of (ER1), the only relevant Christoffel symbols are Γ^0_{01}, Γ^1_{00}, and Γ^1_{11}. Thus,

$$\begin{aligned}
R_{00} &= -\Gamma^1_{00,1} + \Gamma^1_{00}\Gamma^0_{01} - \Gamma^1_{00}\Gamma^1_{11} \\
R_{11} &= \Gamma^0_{01,1} + \Gamma^0_{10}\Gamma^0_{10} - \Gamma^1_{11}\Gamma^0_{01} \\
R_{22} &= 0 \\
R_{33} &= 0
\end{aligned}$$

Now,

$$\begin{aligned}
\Gamma^1_{00} &= -\frac{1}{2}g^{11}g_{00,1} \\
\Gamma^0_{01} &= \frac{1}{2}g^{00}g_{00,1} \\
\Gamma^1_{11} &= \frac{1}{2}g^{11}g_{11,1}
\end{aligned}$$

so

$$\begin{aligned}
R_{00} &= -\frac{1}{2}(g^{11}(-g_{00,1}))_{,1} - \left(\frac{1}{2}(-g_{00})g_{00,1}\right)\left(\frac{1}{2}\frac{1}{g_{00}}g_{00,1}\right) + \left(\frac{1}{2}(-g_{00})g_{00,1}\right) \\
&= -\frac{1}{2}g_{00}g_{00,1,1}\ .
\end{aligned}$$

and

$$\begin{aligned}
R_{11} &= \frac{1}{2}\frac{g_{00,1,1}}{g_{00}} - \frac{1}{2}\left(\frac{g_{00,1}}{g_{00}}\right)^2 + \frac{1}{2}\left(\frac{g_{00,1}}{g_{00}}\right)^2 \\
&= \frac{1}{2}\frac{g_{00,1,1}}{g_{00}}\ .
\end{aligned}$$

Thus, $R_{11} = 0$ means that $g_{00,1,1} = 0$. $R_{00} = 0$ and $g_{11} = 0$ mean the same thing.

If $g_{00} = y(x^1)$, then $y'' = 0$, or

$$\Delta y = 0$$

which is the Laplace equation for y. Therefore, y could be any function of x^1 (let us call it r) such that $\Delta y = 0$. So g_{00} may be a combination of a constant and $\frac{1}{r}$, such as the following:

$$g_{00} = A + \frac{B}{r}$$

with A and B being constants.

0.9.4 Calculation of the Christoffel Symbols Under Some Simplifications

Assume

1. A diagonal matrix.
2. Static conditions: $g_{\mu\nu,0} = 0$.

Then, since $g_{\mu\nu} = \dfrac{1}{g^{\mu\nu}}$ and $g_{\mu\nu} = 0$ for $\mu \neq \nu$,

$$\Gamma^{\alpha}_{\mu\nu} = g^{\alpha\alpha}\Gamma_{\mu\alpha\nu} = g^{\alpha\alpha}\frac{1}{2}(g_{\mu\alpha,\nu} + g_{\alpha\nu,\mu} - g_{\mu\nu,\alpha}) \ .$$

1. For $\alpha = 0$:
$$\Gamma^{0}_{\mu\nu} = g^{00}\frac{1}{2}(g_{\mu 0,\nu} + g_{0\nu,\mu}) \ .$$
Then, only $\mu = 0$ and/or $\nu = 0$ will make $\Gamma^{0}_{..} \neq 0$. Thus, only $\Gamma^{0}_{00}, \Gamma^{0}_{01}, \Gamma^{0}_{02}, \Gamma^{0}_{03}$ may be different from zero:
$$\Gamma^{0}_{00} = \frac{1}{2}g^{00}(g_{00,0} + g_{00,0}) = 0 \qquad \Gamma^{0}_{01} = \frac{1}{2}g^{00}g_{00,1}$$
$$\Gamma^{0}_{02} = \frac{1}{2}g^{00}g_{00,2} \qquad \Gamma^{0}_{03} = \frac{1}{2}g^{00}g_{00,3}$$

2. For $\alpha = 1$:
$$\Gamma^{1}_{\mu\nu} = g^{11}\frac{1}{2}(g_{\mu 1,\nu} + g_{1\nu,\mu} - g_{\mu\nu,1}) \ .$$
The first two terms are zero if $\mu \neq 1$ or $\nu \neq 1$, and the last term is zero if $\nu \neq \mu$. The values of μ or ν are 0, 1, 2, or 3. The couples (μ, ν) are (0,0), (0,1), (0,2), (0,3), (1,1), (1,2), (1,3), (2,2), (2,3), (3,3). Thus:
$$\Gamma^{1}_{00} = \frac{1}{2}g^{11}(-g_{00,1}) \qquad \Gamma^{1}_{01} = 0 \qquad \Gamma^{1}_{02} = 0$$
$$\Gamma^{1}_{03} = 0 \qquad \Gamma^{1}_{11} = \frac{1}{2}g^{11}g_{11,1} \qquad \Gamma^{1}_{12} = \frac{1}{2}g^{11}g_{11,2}$$
$$\Gamma^{1}_{13} = \frac{1}{2}g^{11}g_{11,3} \qquad \Gamma^{1}_{22} = \frac{1}{2}g^{11}(-g_{22,1}) \qquad \Gamma^{1}_{23} = 0$$
$$\Gamma^{1}_{33} = \frac{1}{2}g^{11}(-g_{33,1})$$

3. For $\alpha = 2$:

$$\Gamma^2_{\mu\nu} = \frac{1}{2} g^{22} (g_{\mu 2,\nu} + g_{2\nu,\mu} - g_{\mu\nu,2}) \ .$$

Thus

$$\begin{array}{lll}
\Gamma^2_{00} = \frac{1}{2} g^{22}(-g_{00,2}) & \Gamma^2_{01} = 0 & \Gamma^2_{02} = 0 \\
\Gamma^2_{03} = 0 & \Gamma^2_{11} = \frac{1}{2} g^{22}(-g_{11,2}) & \Gamma^2_{12} = \frac{1}{2} g^{22} g_{22,1} \\
\Gamma^2_{13} = 0 & \Gamma^2_{22} = \frac{1}{2} g^{22} g_{22,2} & \Gamma^2_{23} = \frac{1}{2} g^{22} g_{22,3} \\
\Gamma^2_{33} = \frac{1}{2} g^{22}(-g_{33,2}) & &
\end{array}$$

4. For $\alpha = 3$:

$$\Gamma^3_{\mu\nu} = \frac{1}{2} g^{33} (g_{\mu 3,\nu} + g_{3\nu,\mu} - g_{\mu\nu,3}) \ .$$

Thus

$$\begin{array}{lll}
\Gamma^3_{00} = \frac{1}{2} g^{33}(-g_{00,3}) & \Gamma^3_{01} = 0 & \Gamma^3_{02} = 0 \\
\Gamma^3_{03} = \frac{1}{2} g^{33} g_{33,0} & \Gamma^3_{11} = \frac{1}{2} g^{33}(-g_{11,3}) & \Gamma^3_{12} = 0 \\
\Gamma^3_{13} = \frac{1}{2} g^{33} g_{33,1} & \Gamma^3_{22} = \frac{1}{2} g^{33}(-g_{22,3}) & \Gamma^3_{23} = \frac{1}{2} g^{33} g_{33,2} \\
\Gamma^3_{33} = \frac{1}{2} g^{33} g_{33,3} & &
\end{array}$$

In the previous expressions of the Christoffel symbols, we assume that each matrix element $g_{\mu\mu}$ is a function of x^1, x^2, and x^3. For the rest of this lecture we will assume that g_{00} and g_{11} depend only on x^1, and that g_{22} and g_{33} may depend on x^1 and x^2 but not on x^3. With this in mind, the only $\Gamma^{\cdot}_{\cdot\cdot}$ which are nonzero are:

$$\begin{array}{lll}
\Gamma^0_{01} = \frac{1}{2} g^{00} g_{00,1} & \Gamma^1_{00} = \frac{1}{2} g^{11}(-g_{00,1}) & \Gamma^1_{11} = \frac{1}{2} g^{11} g_{11,1} \\
\Gamma^1_{22} = \frac{1}{2} g^{11}(-g_{22,1}) & \Gamma^1_{33} = \frac{1}{2} g^{11}(-g_{33,1}) & \Gamma^2_{12} = \frac{1}{2} g^{22} g_{22,1} \\
\Gamma^2_{33} = \frac{1}{2} g^{22}(-g_{33,2}) & \Gamma^3_{13} = \frac{1}{2} g^{33} g_{33,1} & \Gamma^3_{23} = \frac{1}{2} g^{33} g_{33,2}
\end{array}$$

The Schwarzchild Metric Tensor

The coordinates used are spherical ones ($x^1 = r$, $x^2 = \theta$, $x^3 = \psi$). Thus the space-time invariant $\mathrm{d}s^2$ has the following form:

$$\mathrm{d}s^2 = c^2 g_{00} \mathrm{d}t^2 - g_{rr}(\mathrm{d}r)^2 - r^2(\mathrm{d}\Omega)^2$$

where $(\mathrm{d}\Omega)^2$ is the differential solid angle, $(\mathrm{d}\Omega)^2 = (\mathrm{d}\theta)^2 + (\sin\theta)^2(\mathrm{d}\psi)^2$. [redacted] Schwarzchild took g_{00} and g_{11} as functions of r only:

$$\begin{aligned} g_{00} &= e^{2\nu} & g^{00} &= e^{-2\nu} \\ g_{11} &= -e^{2\lambda} & g^{11} &= -e^{-2\lambda} \\ g_{22} &= -r^2 & g^{22} &= -\frac{1}{r^2} \\ g_{33} &= -r^2 \sin^2\theta & g^{33} &= -\frac{1}{r^2 \sin^2\theta} \end{aligned}$$

where ν and λ are smooth functions of r. In other words, the matrix is

$$\begin{pmatrix} e^{2\nu} & 0 & 0 & 0 \\ 0 & -e^{2\lambda} & 0 & 0 \\ 0 & 0 & -r^2 & 0 \\ 0 & 0 & 0 & -r^2 \sin^2\theta \end{pmatrix}$$

Also,

$$\begin{aligned} g_{00,1} &= 2\nu' e^{2\nu} & g_{11,1} &= -2\lambda' e^{2\lambda} \\ g_{22,1} &= -2r & g_{33,1} &= -2r \sin^2\theta \\ g_{33,2} &= -2r^2 \sin\theta\cos\theta \end{aligned}$$

Then the Christoffel symbols different from zero are

$$
\begin{aligned}
\Gamma^0_{01} &= \frac{1}{2} g^{00} g_{00,1} = \frac{1}{2} e^{-2\nu}(2\nu' e^{2\nu}) = \nu' \\
\Gamma^1_{00} &= -\frac{1}{2} g^{11} g_{00,1} = -\frac{1}{2}(-e^{-2\lambda})(2\nu' e^{2\nu}) = \nu' e^{2(\nu-\lambda)} \\
\Gamma^1_{11} &= \frac{1}{2} g^{11} g_{11,1} = \frac{1}{2}(-e^{-2\lambda})(-2\lambda' e^{2\lambda}) = \lambda' \\
\Gamma^1_{22} &= -\frac{1}{2} g^{11} g_{22,1} = -\frac{1}{2}(-e^{-2\lambda})(-2r) = -re^{-2\lambda} \\
\Gamma^2_{12} &= \frac{1}{2} g^{22} g_{22,1} = \frac{1}{2}(-\frac{1}{r^2})(-2r) = \frac{1}{r} \\
\Gamma^2_{33} &= -\frac{1}{2} g^{22} g_{33,2} = -\frac{1}{2}(-\frac{1}{r^2})(-2r^2 \sin\theta\cos\theta) = -\sin\theta\cos\theta \\
\Gamma^1_{33} &= -\frac{1}{2} g^{11} g_{33,1} = -\frac{1}{2}(-e^{-2\lambda})(-2r\sin^2\theta) = -re^{-2\lambda}\sin^2\theta \\
\Gamma^3_{13} &= \frac{1}{2} g^{33} g_{33,1} = \frac{1}{2}(-\frac{1}{r^2\sin^2\theta})(-2r\sin^2\theta) = \frac{1}{r} \\
\Gamma^3_{23} &= \frac{1}{2} g^{33} g_{33,2} = \frac{1}{2}(-\frac{1}{r^2\sin^2\theta})(-2r^2\sin\theta\cos\theta) = \cot\theta
\end{aligned}
$$

Now, let us write the Einstein relationship using the Schwarzchild matrix:

$$
\begin{aligned}
R_{00} &= (-\nu'' - \nu'^2 + \nu'\lambda' - \frac{2\nu'}{r})e^{2(\nu-\lambda)} = 0 \\
R_{11} &= \nu'' + \nu'^2 - \lambda'\nu' - \frac{2\lambda'}{r} = 0 \\
R_{22} &= -1 + e^{-2\lambda}(1 + r\nu' - r\lambda') = 0 \\
R_{33} &= R_{22}\sin^2\theta
\end{aligned}
$$

For $R_{00} = 0$ and $R_{11} = 0$, we get

$$\nu' + \lambda' = 0$$

or

$$\nu + \lambda = \text{const.}$$

Assuming a flat space for $r \to \infty$, the constant must be zero. Thus,

$$\nu = -\lambda \ .$$

Now, eliminating λ in $R_{22} = 0$, we get

$$-1 + e^{2\nu}(1 + 2r\nu') = 0$$

or

$$\frac{\mathrm{d}}{\mathrm{d}r}(re^{2\nu}) = 1$$

or

$$re^{2\nu} = r + \text{const.}$$

Let us call the constant $-2m$. Thus,

$$re^{2\nu} = r - 2m \ .$$

Therefore,

$$g_{00} = e^{2\nu} = 1 - \frac{2m}{r} \ . \tag{0.9.32}$$

The same form that we obtain with the coordinates x^0 and x^1, $-\frac{2m}{r}$ is the potential due to gravity which deviates g_{00} from 1 (flat space). Now,

$$g_{11} = -e^{2\lambda} = -e^{-2\nu} = -\frac{1}{g_{00}} = -\frac{1}{1 - \frac{2m}{r}} \ .$$

Let us show now that this solution which establishes that $\nu' = -\lambda'$ and $e^{2\nu} = 1 - \frac{2m}{r}$ satisfies the Einstein relationship $R_{00} = 0$. For that, we must have

$$\nu'' + \nu'^2 - \lambda'\nu' + \frac{2\nu'}{r} = 0$$

or

$$\nu'' + 2\nu'^2 + \frac{2\nu'}{r} = 0 \ . \tag{0.9.33}$$

From (0.9.32),

$$2\nu = \ln(1 - \frac{2m}{r})$$

or

$$\nu' = \frac{m}{r^2 - 2mr}$$

and

$$\nu'' = -\frac{2m(r-m)}{(r^2 - 2mr)^2} \ .$$

Substituting the above into (0.9.33),

$$\frac{-2m(r-m)}{(r^2-2mr)^2} + \frac{2m^2}{(r^2-2mr)^2} + \frac{2m}{(r^2-2mr)r} \ .$$

Multiplying and dividing the last term by $(r^2 - 2mr)$, we find that the above is equal to zero. Thus, $R_{00} = 0$, and $e^{2\nu} = 1 - \dfrac{2m}{r}$ makes all Ricci tensors equal to zero. This is condition imposed onto the fourth dimensional curvature by Einstein.

With the metric tensor

$$\begin{pmatrix} 1 - \dfrac{2m}{r} & 0 & 0 & 0 \\ 0 & -\dfrac{1}{1-\frac{2m}{r}} & 0 & 0 \\ 0 & 0 & -r^2 & 0 \\ 0 & 0 & 0 & -r^2\sin^2\theta \end{pmatrix}$$

the fourth dimension invariant length, after using the Einstein relationship (i.e., the Ricci tensor equal to zero), becomes

$$\mathrm{d}s^2 = (1 - \frac{2m}{r})c^2\mathrm{d}t^2 - \frac{\mathrm{d}r^2}{1 - \dfrac{2m}{r}} - r^2\mathrm{d}\theta^2 - r^2\sin^2\theta\ \mathrm{d}\psi^2 \ .$$

By using the variables x^2 and x^3, such that $\mathrm{d}x^2 = r\mathrm{d}\theta$ and $\mathrm{d}x^3 = r\sin\theta\ \mathrm{d}\psi$, then

$$\mathrm{d}s^2 = (1 - \frac{2m}{r})c^2\mathrm{d}t^2 - \frac{\mathrm{d}r^2}{1 - \dfrac{2m}{r}} - (\mathrm{d}x^2)^2 - (\mathrm{d}x^3)^2 \ .$$

The metric tensor then becomes

$$\begin{pmatrix} 1-\dfrac{2m}{r} & 0 & 0 & 0 \\ 0 & -\dfrac{1}{1-\frac{2m}{r}} & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix}$$

and the determinant $|g| = -1$, just like in the Lorentz matrix.

0.9.5 The Hubble Law

From observed red shifts, in 1928 Hubble found that the galaxies recede at a velocity proportional to the distance,

$$v = Hr \ . \tag{0.9.34}$$

Let us assume that this law is valid for all times t, or

$$v(t) = H(t)r(t)$$

or

$$\frac{\mathrm{d}r(t)}{\mathrm{d}t} = H(t)r(t) \ .$$

Integrating,

$$\ln\frac{r}{r_0} = \int_{t_0}^{t} H(t)\ \mathrm{d}t$$

or

$$\frac{r}{r_0} = e^{\int_{t_0}^{t} H(t)\ \mathrm{d}t} \ . \tag{0.9.35}$$

The ratio $\frac{r(t)}{r_0(t_0)}$ is called $\bar{R}(t)$. As time passes, the galaxy's distance increase by this factor $\bar{R}(t)$, which is related to the Hubble parameter H by

$$r(t) = \bar{R}(t)r_0(t_0) = r_0(t_0)e^{\int_{t_0}^{t} H(t)\ \mathrm{d}t}$$

or

$$H(t) = \frac{\mathrm{d}}{\mathrm{d}t}[\ln(\bar{R}(t))] \ . \tag{0.9.36}$$

The function $H(t)$ may take on different forms depending on what parameters are used to define the universe. Next, we will discuss particular cases of H as a time-dependent parameter.

1. $H(t)$ is constant and independent of time, so

$$v(t) = \frac{\mathrm{d}r}{\mathrm{d}t} = Hr(t) \tag{0.9.37}$$

or by integrating,

$$\ln\frac{r}{r_0} = H\int_{t_0}^{t} \mathrm{d}t = H(t-t_0)$$

or

$$r(t) = r_0(t_0)e^{H(t-t_0)} \ . \tag{0.9.38}$$

The acceleration of the galaxy will be

$$\alpha(t) = \frac{\mathrm{d}v(t)}{\mathrm{d}t} = H\frac{\mathrm{d}r(t)}{\mathrm{d}t} = Hv(t) \ . \tag{0.9.39}$$

or

$$\ln\left(\frac{v(t)}{v_0(t_0)}\right) = \int_{t_0}^{t} H \ \mathrm{d}t$$

or

$$v(t) = v_0(t_0)e^{H(t-t_0)} \ .$$

Therefore,

$$\alpha(t) = Hv_0(t_0)e^{H(t-t_0)} \ . \tag{0.9.40}$$

Also, from (0.9.39),

$$\alpha(t) = Hv(t) = H(Hr(t)) = H^2r(t) \ .$$

Thus,

$$\alpha(t) = H^2r(t) \ . \tag{0.9.41}$$

A force $F(t)$ on the unit mass is

$$F(t) = \alpha(t) = H^2r(t)$$

or a potential U such that $F = -\nabla_r U$ is

$$U(t) = -H^2\frac{r^2}{2} \ . \tag{0.9.42}$$

2. Let us assume that

$$H(t) = \frac{1}{t} \ .$$

The Hubble Law says that $v = rH$, but $v = \dfrac{\mathrm{d}r}{\mathrm{d}t}$, so

$$v(t) = \frac{\mathrm{d}r}{\mathrm{d}t} = r(t)H \ .$$

Integrating,

$$\ln\frac{r}{r_0} = \ln\frac{t}{t_0}$$

or

$$\bar{R} = \frac{r}{r_0} = \frac{t}{t_0} \ .$$

Thus,

$$\frac{v(t)}{v_0(t_0)} = \frac{r(t)t_0}{r_0(t_0)t} = 1 \tag{0.9.43}$$

which means that $v = v_0$, or v is constant and independent of time. In this case, the acceleration $\alpha = 0$ and there are no net forces acting in the galaxy. We observe that as $t \to 0$, $H \to \infty$, which corresponds to the initial "big bang" where all the distances $r = r_0$ were zero. Also, as $t = t_0$, $\bar{R} = 1$. Since $\bar{R} = \dfrac{r}{r_0}$, the number $N(t)$ density which is proportioanl to r^3 are proportional to t^3. That is,

$$\frac{N(t)}{N_0(t_0)} = \left(\frac{t}{t_0}\right)^3 \ . \tag{0.9.44}$$

3. Let us assume that

$$H(t) = \frac{\beta}{t} \ . \tag{0.9.45}$$

The Hubble Law says

$$v(t) = r(t)H = \beta\frac{r(t)}{t} \ . \tag{0.9.46}$$

Since $v(t) = \dfrac{\mathrm{d}r(t)}{\mathrm{d}t}$, then by integrating we get

$$\ln\frac{r(t)}{r_0(t_0)} = \ln\left(\frac{t}{t_0}\right)^\beta \ .$$

Thus,

$$\bar{R}(t) = \frac{r(t)}{r_0(t_0)} = \left(\frac{t}{t_0}\right)^{\beta} . \tag{0.9.47}$$

Now, from (0.9.46)

$$v(t) = \beta \frac{r(t)}{t} = \beta \frac{\bar{R}(t)}{t} r_0(t_0) = \frac{v_0(t_0)}{t}\left(\frac{t}{t_0}\right)^{\beta} t_0 = v_0(t_0)\left(\frac{t}{t_0}\right)^{\beta-1} \tag{0.9.48}$$

since

$$v_0(t_0) = \beta \frac{r_0(t_0)}{t_0} . \tag{0.9.49}$$

From (0.9.48)

$$\frac{v(t)}{v_0(t_0)} = \left(\frac{t}{t_0}\right)^{\beta-1} .$$

For $\beta = 1$, we have the same result as before: v is a constant. For $\beta \neq 1$,

$$\alpha = \frac{\mathrm{d}v(t)}{\mathrm{d}t} = \frac{\mathrm{d}}{\mathrm{d}t}\left(\left(\frac{t}{t_0}\right)^{\beta-1}\right) v_0(t_0) = (\beta-1)\frac{t^{\beta-2}}{t_0^{\beta-1}} v_0(t_0) . \tag{0.9.50}$$

Let us substitute v_0 by r_0 using equation (0.9.49):

$$\alpha = \frac{(\beta-1)}{t_0^{\beta-1}} \frac{r_0\beta}{t_0} t^{\beta-2} = \left(\beta(\beta-1)\frac{r_0(t_0)}{t_0^{\beta}}\right) t^{\beta-2} . \tag{0.9.51}$$

Again for $\beta = 1$ $\alpha = 0$.

Let us get the acceleration (which is the force per unit mass) as a function of $r(t)$, assuming r_0 and t_0 as given values. Since

$$\frac{r(t)}{r_0(t_0)} = \left(\frac{t}{t_0}\right)^{\beta}$$

or simply

$$\frac{r}{r_0} = \left(\frac{t}{t_0}\right)^{\beta} .$$

Thus

$$t = \left(\frac{r}{r_0}\right)^{\frac{1}{\beta}} t_0$$

and

$$t^{\beta-2} = \left[\left(\frac{r}{r_0}\right)^{\frac{1}{\beta}} t_0\right]^{\beta-2} = \left(\frac{r}{r_0}\right)^{\frac{\beta-2}{\beta}} t_0^{\beta-2} \ . \tag{0.9.52}$$

Using equation (0.9.52), the acceleration becomes

$$\alpha = \beta(\beta-1)\frac{r_0}{t_0^{\beta}}\left(\frac{r}{r_0}\right)^{\frac{\beta-2}{\beta}} t_0^{\beta-2} = \beta(\beta-1)t_0^{-2} r_0^{\frac{2}{\beta}} r^{\frac{\beta-2}{\beta}} \ .$$

In other words,

$$\alpha = \xi r^{\frac{\beta-2}{\beta}} \tag{0.9.53}$$

where

$$\xi = \beta(\beta-1) r_0^{\frac{2}{\beta}} t_0^{-2} \ . \tag{0.9.54}$$

For $\beta = \dfrac{2}{3}$, we have a force inversely proportional to r^2, which is a Newtonian type of force.

0.9.6 Expansion of the Universe Under the Fried- [redacted] mann Fundamental Metric and a Tensor Force Acting on the Material of the Galaxy [redacted]

In a simple fluid dynamic, classically the balance of matter in a given element of volume dV is given by the continuity equation

$$\frac{\partial \rho}{\partial t} + \nabla \cdot \vec{j} = 0 \tag{0.9.55}$$

where $\vec{j}$ is the current density associated with the velocity by $\vec{j} = \rho \vec{u}$. If the element of volume suffers stresses from the rest of the continuous media, then the equation much change to include those internal forced plus the possibility of external forces.

In four dimensions, we define the energy-current density J^{μ}. Observe that we have added the word "energy" to the current density, because after the equivalence between mass and energy we want to include those terms which classically may be studied only as energy contributions different form the mass-related terms. By definition,

$$J^{\mu} = \rho_0 U^{\mu} = \rho_0 \gamma (c, \vec{u}) = (\rho c, \vec{j}) \tag{0.9.56}$$

where

$$\vec{j} = \rho \vec{u} \quad \text{and} \quad \rho = \rho_0 \gamma \ . \tag{0.9.57}$$

γ accounts for the length contraction when going from system S to the S_0 or proper system (see Lecture 7).

The new equation of continuity in four dimensions is the following:

$$J^{\mu}{}_{,\nu} = 0 \tag{0.9.58}$$

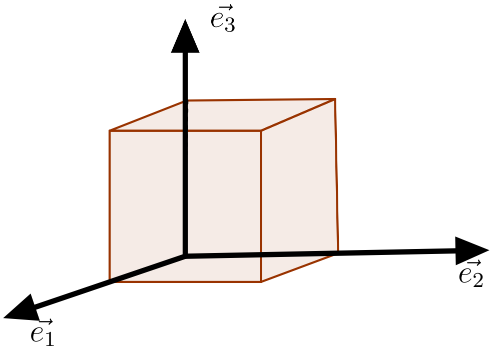
where the coordinate variables are: $\nu = 0$, $x^0 = ct$; $\nu = 1$, $x^1 = x$; $\nu = 2$, $x^2 = y$; $\nu = 3$, $x^3 = z$.

This equation of continuity is not complete because our test element of volume may be affected by stresses acting on the element of volume dV through the walls of the element dV.

In analogy to the classic studies of continuous media, we assume that the following tensor describes those actions in the rest frame S_0:

$$T_0^{\mu\nu} = \begin{pmatrix} c^2\rho_0 & 0 & 0 & 0 \\ 0 & t_0^{11} & t_0^{12} & t_0^{13} \\ 0 & t_0^{21} & t_0^{22} & t_0^{23} \\ 0 & t_0^{31} & t_0^{32} & t_0^{33} \end{pmatrix}$$

The tensor T_0^{ij} may be understood in the following way. Assume a surface enclosing a differential unit of cubic volume $\mathrm{d}V$ of the fluid within the three orthogonal axes (x, y, z). The stress to the fluid on each of its sides is described by the two index tensor t^{ij} where j stands for the axis which is normal to the face of the cube and i the component of the stress on that surface in the direction of the axis i. The unit vectors are $\vec{e_1}$, $\vec{e_2}$, $\vec{e_3}$.



Let us apply the Gauss divergence theorem to the cube. The force on the face $\mathrm{d}s_j$ perpendicular to $\vec{e_j}$ is $t^{ij}\ \mathrm{d}s_j$. The flux on the forces with direction of $\vec{e_i}$ is $\int t^{ij}\ \mathrm{d}s_j$, which is equal to $\int t^{ij}_{,j}\ \mathrm{d}V$ where $j = 1, 2, 3$. Thus, by the theorem we identify the force $\vec{f^i}$ due to the stress as the divergence of the tensor t^{ij}:

$$\vec{f^i} = t^{ij}_{,j}\ .$$

Also, the tensor expressed in the rest frame as t_0^{ij} may be transformed to the laboratory frame as

$$\begin{pmatrix} t^{11} & t^{12} & t^{13} \\ t^{21} & t^{22} & t^{23} \\ t^{31} & t^{32} & t^{33} \end{pmatrix}$$

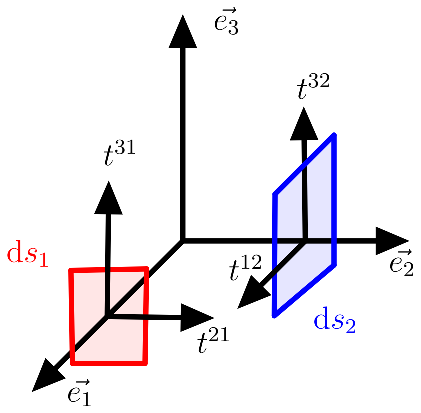
where $t^{ii} = t_0^{ii}$, $t_{i<j}^{ij} = t_0^{ij}\gamma$, and $t_{i>j}^{ij} = \dfrac{t_0^{ij}}{\gamma}$. The expression for γ is the following:

$$\gamma(u) = \frac{1}{\sqrt{1 - \dfrac{u^2}{c^2}}}$$

where u is the velocity of the fluid in the direction of $\vec{e_1}$. The above transformation may be derived from the fact (see *Introduction to Special Relativity* by Wolfgang Rindler) that the force $\vec{f}$ on the element of area $\mathrm{d}s_j$, $\vec{f} = t^{ij}\ \mathrm{d}s_j$, is equal in both systems S and S_0 when the velocity u is parallel to $\vec{f}$ and the mass m_0 measured in S_0 remains constant. Also, we must take into account that when $\mathrm{d}s_j$ has a side parallel to $\vec{e_1}$ (or $\vec{u_1}$) then as it moves $\mathrm{d}s_j$ will suffer a contraction. For instance, $\mathrm{d}s_2 = \Delta x \cdot \Delta z$, so $\mathrm{d}s_{02} = \mathrm{d}s_2\gamma(u)$. Thus, $\mathrm{d}s_2 t^{12} = \mathrm{d}s_{02} t_0^{12} = t_0^{12}\gamma(u)\mathrm{d}s_2$, and therefore $t^{12} = t_0^{12}\gamma(u)$.

For the other two faces perpendicular to $\vec{e_1}$ we have $\mathrm{d}s = \mathrm{d}s_0$, as the motion of them will not have any relativistic effect. For example, the force on the face perpendicular to $\vec{e_1}$ in the direction of $\vec{e_2}$ is $t^{21}\mathrm{d}s_1 = \gamma(u)(1 - \dfrac{uv}{c^2})t_0^{21}\mathrm{d}s_{01} = \dfrac{\gamma(u')}{\gamma(u)}t_0^{21}\mathrm{d}s_{01}$. In general, $\dfrac{\gamma(u')}{\gamma(u)} = \gamma(u)(1 - \dfrac{uv}{c^2})$. In our case, u' (the velocity with respect to the rest frame) is zero, so $\gamma(u') = 1$ and $t^{21} = \dfrac{t_0^{21}}{\gamma(u)}$, since $\mathrm{d}s_1 = \mathrm{d}s_{01}$ (no length contraction).

For the diagonal element of the stress tensor like t^{11} we have no variation in the area nor in the force. In fact, according to the transformation of the forces when the force is parallel to $\vec{u}$, $t_0^{11}\mathrm{d}s_{01} = t^{11}\mathrm{d}s_1$. But as the surface moves with the fluid, $\mathrm{d}s_{01} = \mathrm{d}s_0$ (no contraction). Thus, $t_0^{11} = t^{11}$.



By changing to the frame S, respect to which S_0 moves with velocity $\vec{u}$, this tensor should include the energy due to motion. We may write it as

$$T^{\mu\nu} = \begin{pmatrix} c^2\rho & cg^1 & cg^2 & cg^3 \\ cg^1 & p^{11} & p^{12} & p^{13} \\ cg^2 & p^{21} & p^{22} & p^{23} \\ cg^3 & p^{31} & p^{32} & p^{33} \end{pmatrix}$$

The energy density is ρc^2, the total momentum density or energy current is cg^i, and the total stress on the unit area as seen from the laboratory system is p^{ij}.

Let us see next how g^i and P^{ij} are related to the original tensor t^{ij}. Let us assume that f^i is the external force on $\mathrm{d}V$ which is contrasted by the stress t^{ij} which creates a force in the i direction equal to $t^{ij}_{,j}$ per unit volume. Thus, the equation of motion is

$$\frac{\mathrm{d}}{\mathrm{d}t}(g^i \mathrm{d}V) = (f^i - t^{ij}_{,j})\mathrm{d}V \ . \tag{0.9.59}$$

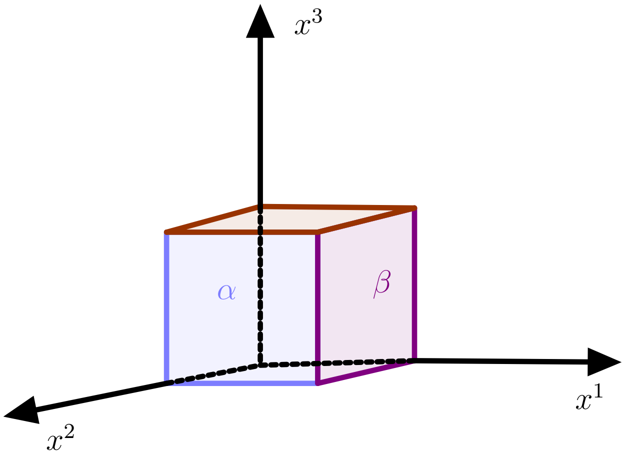
The first side the equation implies the variation of g^i, as well as the variation of $\mathrm{d}V$.

In order to see how t^{ji} is part of g^i let us imagine a surface through which the fluid moves across with velocity $\vec{u}$. The energy transmitted across in the direction of $\vec{e_i}$ is $t^{ji}\vec{e_j} \cdot \vec{u} = t^{ji}u_j$, where repeated indexes implies summation. This power is transformed into momentum density by dividing by c^2 in accordance with the relationship $m = \dfrac{E}{c^2}$. So the momentum density is $\dfrac{t^{ji}u_j}{c^2}$. As we have said before, the total momentum density will then be

$$g^i = \rho u^i + u_j \frac{t^{ji}}{c^2} \ . \tag{0.9.60}$$

If heat or other types of energy had been crossing the imaginary unit plane, they should have been added to the previous expression of the momentum density. As we will see, all this extra energy will never have shown up in the equation if it were not for the equivalence between energy and mass. Observe that this extra energy is crossing the area instantaneously, so that the motion of the fluid is unimportant. The motion is taken into account in the first term, ρu^i.

Returning to the equation of motion (0.9.59), we observe that the element of volume will change with time. In fact, in the laboratory frame with coordinates (ct, x^1, x^2, x^2), the face α (see figure) in an interval dt sweeps a volume proportional to $u^1 \mathrm{d}x^2 \mathrm{d}x^3 \mathrm{d}t$, while the face β sweeps a volume of $\left(u^1 + \dfrac{\partial u^1}{\partial x^1}\mathrm{d}x^1\right)\mathrm{d}x^2 \mathrm{d}x^3 \mathrm{d}t$.



Thus, the change in volume in dt is

$$\frac{\mathrm{d}V}{\mathrm{d}t} = \frac{\partial u^1}{\partial x^1}\mathrm{d}V \ . \tag{0.9.61}$$

So equation (0.9.59) becomes

$$\mathrm{d}V\frac{\mathrm{d}}{\mathrm{d}t}g^i + g^i\frac{\partial u^j}{\partial x^j}\mathrm{d}V = (f^i - t^{ij}_{,j})\mathrm{d}V \ . \tag{0.9.62}$$

The first term may be explicitly written as

$$\frac{\mathrm{d}}{\mathrm{d}t}g^i = \frac{\partial g^i}{\partial t} + \frac{\partial g^i}{\partial x^j}\frac{\mathrm{d}x^j}{\mathrm{d}t} = \frac{\partial g^i}{\partial t} + \frac{\partial g^i}{\partial x^j}u^j \ . \tag{0.9.63}$$

Then (0.9.62) becomes

$$\frac{\partial g^i}{\partial t} + \frac{\partial g^i}{\partial x^j}u^j + g^i\frac{\partial u^j}{\partial x^j} = f^i - t^{ij}_{,j}$$

or

$$\frac{\mathrm{d}g^i}{\mathrm{d}t} + \frac{\partial}{\partial x^j}(g^iu^j + t^{ij}) = f^i \ . \tag{0.9.64}$$

The four-dimensional current which is given by

$$J = \rho_0 U = \rho_0\gamma(c,u) = (c\rho, \rho\vec{u}) = (c\rho, \vec{j})$$

has a divergence of zero. That is,

$$J^{\mu}_{,\mu} = 0 \ . \tag{0.9.65}$$

This expresses the continuity of matter in four dimensions, similar to the classic continuity equation (0.9.55).

The four-dimensional force is defined as

$$F_0^{\mu} = m_0 \frac{\mathrm{d}U}{\mathrm{d}\tau} = \gamma \frac{\mathrm{d}}{\mathrm{d}t}(m_0 U) = \gamma \frac{\mathrm{d}}{\mathrm{d}t}(\gamma m_0 (c, \vec{u})) = \gamma \left(\frac{1}{c} \frac{\mathrm{d}E}{\mathrm{d}t}, \vec{f} \right) \ . \tag{0.9.66}$$

In the case of a pure force $(\frac{\mathrm{d}m_0}{\mathrm{d}t} = 0)$,

$$\frac{\mathrm{d}E}{\mathrm{d}t} = (\vec{f} \cdot \vec{u})$$

and so

$$F_0^{\mu} = \gamma \left(\frac{\vec{f} \cdot \vec{u}}{c}, \vec{f} \right) \tag{0.9.67}$$

When an external force is acting on the element of volume of unit value, then the force F_0^{μ} is related to the tensor $T^{\mu\nu}$ by the equation of continuity

$$T^{\mu\nu}_{,\nu} = F_0^{\mu} \ . \tag{0.9.68}$$

For $\mu = 0$, we have $T^{0\nu} = (c^2\rho, cg^1, cg^2, cg^3)$. For $\mu = 1$ to 3, we have $T^{i\nu} = (cg^i, p^{i1}, p^{i2}, p^{i3})$. Thus, there are two types of equations as a result of (0.9.68):

$$\frac{\partial c^2 \rho}{\partial t} + \nabla \cdot (c^2 \vec{g}) = \vec{f} \cdot \vec{u} \tag{0.9.69}$$

and

$$\frac{\partial g^i}{\partial t} + \frac{\partial p^{ij}}{\partial x^j} = \vec{f}^i \tag{0.9.70}$$

The equation (0.9.69) implies the balance of energy in the unit volume of the fluid. The first term expresses the energy rate inside the volume, while the divergence of $c^2 g$ taken to the second side of the equation is the energy current density that is received by the fluid through the surface of the element of volume. Finally, $(\vec{f} \cdot \vec{u})$ is the power delivered by the external force of $\vec{f}$ into the element of volume.

Comparing (0.9.70) and (0.9.64), we see that

$$P^{ij} = g^i u^j + t^{ij} \ . \tag{0.9.71}$$

By using (0.9.60), P^{ij} is expressed explicitly as a function of t^{ij}.

A Particular Case of the Tensor Force $T^{\mu\nu}$

Let us assume that, just like in an ideal gas,

$$T^{\mu}_{\nu} = \begin{pmatrix} \rho c^2 & 0 & 0 & 0 \\ 0 & -P & 0 & 0 \\ 0 & 0 & -P & 0 \\ 0 & 0 & 0 & -P \end{pmatrix}$$

It is a diagonal matrix with ρ dependent on time, but P is constant and isotropic. We would like to work with the matrix $T^{\mu\nu}$ and the Friedmann metric tensor $g_{\mu\nu}$.

$$T^{\mu\nu} = T^{\mu}_{\beta} g^{\beta\nu} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -P & 0 & 0 \\ 0 & 0 & -P & 0 \\ 0 & 0 & 0 & -P \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -\frac{1}{a^2} & 0 & 0 \\ 0 & 0 & -\frac{1}{a^2} & 0 \\ 0 & 0 & 0 & -\frac{1}{a^2} \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & \frac{P}{a^2} & 0 & 0 \\ 0 & 0 & \frac{P}{a^2} & 0 \\ 0 & 0 & 0 & \frac{P}{a^2} \end{pmatrix}$$

In Lecture 4 we showed that the Ricci tensor has an absolute (or covariant) derivative equal to zero. We also showed that the Einstein equation under the action of a tensor $T^{\mu\nu}$ is

$$R^{\mu\nu} = -\frac{1}{2} g^{\mu\nu} R = \chi T^{\mu\nu} \tag{0.9.72}$$

where $R^{\mu\nu}$ is the Ricci tensor, and since the absolute derivative of the first side of the equation is zero, then

$$\mathrm{D}_{\nu} T^{\mu\nu} = 0$$

This conservative equation can be written as

$$\mathrm{D}_\lambda T^{\mu\nu} = \frac{\partial T^{\mu\nu}}{\partial x^\lambda} + \Gamma^\nu_{\lambda\alpha} T^{\mu\alpha} + \Gamma^\mu_{\lambda\alpha} T^{\nu\alpha} = 0 \ .$$

Then for $\lambda = \mu$ it becomes

$$\frac{\partial T^{\mu\nu}}{\partial x^\mu} + \Gamma^\nu_{\mu\alpha} T^{\mu\alpha} + \Gamma^\mu_{\mu\alpha} T^{\nu\alpha} = 0 \ . \tag{0.9.73}$$

The following table summarizes the terms different from zero in the previous equation, given that the only non-zero Christoffel symbols for the Friedmann matrix are

$$\Gamma^0_{ii} = \frac{a\dot{a}}{c} \quad \text{and} \quad \Gamma^i_{0i} = \frac{\dot{a}}{ca} \tag{0.9.74}$$

where $i = 1, 2, 3$.

ν	μ	$\lambda = \mu$	α	Non-Zero Term
0	0	0	0	$T^{00}_{,0} = c\dot{\rho}$
0	i	i	i	$\Gamma^\nu_{\alpha\mu} T^{\alpha\mu} = \frac{\dot{a}}{ca} P$
0	i	i	0	$\Gamma^\mu_{\mu\alpha} T^{\alpha\nu} = \frac{\dot{a}}{a} \rho c$

Summing all of these contributors for $\nu = 0$ and $i = 1, 2, 3$, we have the following:

$$c\dot{\rho} + 3\frac{\dot{a}}{ca} P + 3\frac{\dot{a}}{a}\rho c = c^2\dot{\rho} + 3\frac{\dot{a}}{a}(P + \rho c^2) = 0 \ .$$

Particular cases:

1. For $P = 0$, we have
$$\frac{\dot{\rho}}{\rho} = -3\frac{\dot{a}}{a} \ .$$
   Integrating,
$$[\ln \rho]_0^t = -3[\ln a]_0^t \ .$$
   Assuming that the limits for ρ and a for $t = 0$ are $\rho(0) = 1$ and $a(0) = 1$, we have
$$\rho = \frac{1}{a^3} \ . \tag{0.9.75}$$

2. For $P = c^2\frac{\rho}{3}$,

$$c^2\dot{\rho} + 3\frac{\dot{a}}{a}(\frac{4}{3}\rho c^2) = 0$$

and

$$\frac{\dot{\rho}}{\rho} = -4\frac{\dot{a}}{a}\ .$$

Integrating,

$$\ln\rho = \ln(a^{-4})\ .$$

Thus,

$$\rho = \frac{1}{a^4} \tag{0.9.76}$$

with the same assumptions of the limit values of ρ and a for $t = 0$: $\rho(0) = 1$ and $a(0) = 1$.